The Dissertation Committee for Rafaella Sale

certifies that this is the approved version of the following dissertation:

Evaluation of a Progress Monitoring and Feedback Intervention

in a Community Mental Health Center for Youth

**Committee:**

Sarah Kate Bearman, Supervisor

Cindy Carlson

Keenan Pituch

Molly Lopez

# Evaluation of a Progress Monitoring and Feedback Intervention in a Community Mental Health Center for Youth

by

**Rafaella Sale**

**Dissertation**

Presented to Faculty of Graduate School

of the University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

**Doctor of Philosophy**

The University of Texas at Austin

August 2020

# Abstract


**Evaluation of a Progress Monitoring and Feedback Intervention in a Community Mental Health Center for Youth**

by

Rafaella Sale, PhD

The University of Texas at Austin, 2020

SUPERVISOR: Sarah Kate Bearman

Clinicians often have difficulty accurately predicting whether their patients are progressing as planned while enrolled in therapy. Use of a measurement feedback system (MFS) has been shown within the adult literature to improve mental health treatment outcomes, especially for those who are most likely to dropout from treatment. An MFS's effects are theorized to work through clinician behavior (i.e., tailoring treatment) that is elicited in response to receiving patient feedback. The current study examined an MFS's effects on treatment length and dropout rate after being implemented in one community mental health center for youth ($n$ = 538; 57% male, 52% Latinx/Hispanic, M age = 9.80). Level of clinician implementation of the MFS was also examined, and higher levels were found to impact treatment outcomes. Limitations include systemic issues related to archival data, including use of a historical control. Clinical and theoretical implications of these findings are discussed, in addition to recommended next steps for researchers interested in dissemination and implementation of effective mental health treatment.

# Table of Contents

# List of Tables

# List of Figures

Approximately 26-37% of children meet criteria for one or more psychological disorders by adulthood (American Psychiatric Association [APA], 2000; Costello, Mustillo, Erkanli, Keeler, & Angold, 2003; Costello, He, Sampson, Kessler, & Merikangas, 2014). Twenty-five percent of children meet criteria for two psychological disorders at any given time (Costello et al., 2003). Psychotherapeutic intervention is often recommended as the first line of treatment, yet up to 80% of individuals who enroll into mental health services through community providers do not significantly improve over an episode of care (Amble, Gude, Stubdal, Andersen, & Wampold, 2015; Davidson et al., 2017). Approximately half of youth who present for mental health services do not improve and 15-24% worsen while in treatment (Cooper, Stewart, Sparks, & Bunting, 2013; Smith & Jensen-Doss, 2017; Warren, Nelson, Mondragon, Baldwin, & Burlingame, 2010). Clinical worsening, or deterioration, after enrolling into therapy is believed to be grossly underestimated, since individuals must complete measures on at least two separate occasions in order to be documented as "deteriorated" (Boswell, Kraus, Miller, & Lambert, 2015). In actuality, the majority of patients preemptively terminate from treatment (i.e., "dropout") before their symptoms improve (McClintock, Perlman, McCarrick, Anderson, & Himawan, 2017), and never return to complete post-treatment assessment measures.

Therapists typically underestimate their patients' clinical worsening and overestimate their improvement. One study surveyed 129 mental health professionals in private practice to collect their estimates for the proportion of their patients who improve or deteriorate before leaving treatment (Walfish, McAlister, O'Donnell, & Lambert, 2012). Approximately 58% of respondents estimated that 80% or more of their patients improved, and 21% of respondents believed 90% or more of their patients improved before treatment ended. When asked to rate their therapeutic skills compared to other clinicians, 92% of the sample rated themselves above the 75$^{th}$ percentile

(Walfish et al., 2012). It is possible that this study returned responses from a subgroup of highly successful clinicians, though it is more likely that clinicians are poorly equipped to accurately predict patient trajectories throughout care. One study examined clinicians' therapy progress notes in order to evaluate perceptions of deterioration (Hatfield, McCullough, Frantz, & Krieger, 2010). Specifically, researchers reviewed the progress notes of patients who had worsened from one session to the next by at least two times the number of points to count as "reliable change," (i.e., patients who experienced a dramatic deterioration). Even though clinicians had access to their patients' measurement data, less than one-third of their case notes included any information that referenced patients' worsening of symptoms (Hatfield et al., 2010).

Whereas clinical judgment for detecting clinical worsening may be faulty, actuarial methods may be better suited to predict patient trajectories. In another study, 48 clinicians were asked how many patients out of a sample of 550 would end treatment in a worsened state (Hannan et al., 2005). Responses were compared to results from an empirical algorithm that took into account symptom severity at intake and trajectory of response to predict how likely an individual was to drop out from treatment, based on a database of past patient trajectories. Out of the 550 patients, 40 actually deteriorated across sessions before dropping out of treatment. Clinicians successfully predicted deterioration in one of these patients, while the algorithm accurately predicted 36 patients (86% of the deteriorated subsample) (Hannan et al., 2005). Furthermore, the algorithm was able to identify these patients by the 3rd session (Hannan et al., 2005). For youth, a similar actuarial method used to alert clinicians to risk of treatment failure accurately predicted the treatment trajectories for 71% of 2,715 cases, based on youth report alone (Cannon, Warren, Nelson, & Burlingame, 2010).

Clinician subjective judgement has been found to be inconsistent and, at times, biased (Walfish et al., 2012; Tracey, Wampold, Lichtenberg, & Goodyear, 2014). For example, Treichler & Spaulding (2018) provided evidence that type of training and years of experience may impact how a clinician rated patient motivation, social skills, and participation in treatment. The same study also revealed gender bias in play, as males were perceived as being more engaged in treatment than females (Treichler & Spaulding, 2018). In another study, Bar-Kalif and colleagues (2015) compared patient and clinician report on patient's functioning over treatment and found clinicians varied in over- and underestimating their patients' functioning without a consistent pattern. Roughly half of clinicians were more likely to overestimate patient's functioning, while others underestimated with great variability across patients. It was only after clinicians were granted access to their patients' own reports did clinician report begin to align with their patient's (Bar-Kalif et al., 2015). According to scores at the end of treatment, patients who improved the most were seen by clinicians whose reports became more accurate earlier in treatment (Bar-Kalif et al., 2015), suggesting that feedback from patients helps clinicians more accurately gauge patient trajectories and perhaps become more effective providers.

MacDonald and Mellor-Clark (2015) reviewed the social psychology literature and suggest that clinician judgement often falls short because individuals cannot accurately assess their own performance. Expert intuition is more likely to be accurate in "high validity environments," or settings where regularities in behavior are common and expertise is built through experience (p. 251). Statistics and actuarial predictions have been found to outperform experts in irregular environments because of the consistent attention and weight assigned to cues and data. Clinical settings are thought to be irregular, or "low validity environments," where clinician blindsided-ness should be expected and attended to openly (MacDonald & Mellor-Clark, 2015, p. 252). The

expectation that clinicians should be able to remember each patient's state at intake and quickly make comparisons at each time point throughout treatment is unreasonable at best (Hatfield et al., 2010). Acceptance that clinicians are being asked to do the impossible opens the door for the introduction of more precise methods that have been proven to make prediction possible. For instance, the methods embedded into progress monitoring and feedback platforms are empirically-derived and provide expected response curves for each individual patient based on severity at baseline (Lambert et al., 2002a).

The current proposal describes progress monitoring and feedback, a type of intervention that has demonstrated empirical support for improving patient outcomes across a variety of settings. Originally conceptualized as a quality assurance effort, the intervention has been theorized to personalize a clinician's practice to promote patient engagement and prevent treatment failure. Progress monitoring and feedback is delivered through a measurement feedback system, a web-based platform that has the ability to guide treatment planning quickly and efficiently for the clinician. The proposed study involves a program evaluation of a commercialized measurement feedback system that was adopted agency-wide by one large community mental health center. Its impact on premature dropout and treatment length in addition to rate of symptom change was examined. Findings contribute to the implementation science research base through replication of past studies' findings and provision of concrete steps for improving mental health services in real world settings.

Chapter 2

# LITERATURE REVIEW

This chapter begins with an overview of mental health service in the United States and the costs associated with access and utilization of care over time. An approach to improving the practice of psychology follows. The progress monitoring and feedback literature and meta-analyses on the efficacy and effectiveness of measurement feedback systems (MFSs) conducted with adult populations across various mental and behavioral health settings will be reviewed. Next, the handful of studies that have examined MFSs among youth populations will be discussed thoroughly. Methodological issues from adult and child studies will be described in order to contextualize interpretations and the applicability of findings to real world settings. Lastly, the chapter addresses the major theorized mechanisms of change associated with progress monitoring and feedback interventions. A case will be made concerning the utility and cost-effectiveness of MFSs for community mental health settings and the potential value they represent to stakeholders.

## State of Affairs of Mental Health Services in the United States

**SERVICE UTILIZATION.** Over one-third of children and one-fourth of adolescents meet diagnostic criteria for a psychological disorder within any given year (Costello et al., 2003; Costello et al., 2014). The American Psychiatric Association and American Academy of Child and Adolescent Psychiatry have both stated that outpatient psychotherapy should be the first line of treatment for the majority of youth who present with significant psychological distress (Birmaher & Brent, 2007; Connolly & Bernstein, 2007). Sadly, 77% of youth who meet criteria for one or more psychological disorder never access this type of care (Costello et al., 2014). Costello and colleagues (2014) found that within a nationally representative sample of 10,148 families, adolescents were more likely to have received mental health services from their school or general

medical practitioner than a specialist with disorder-specific training. Even for teens with three or more psychiatric diagnoses, less than half had received mental health services from an outpatient mental health facility. Youth were most likely to have accessed care if they were male, had a disruptive behavior disorder, or came from a single parent household classified as low income, to whom mental health services were provided through juvenile justice, alternative school placement, or special education (Costello et al., 2014).

**COSTS OF CARE.** Disorders associated with mental and behavioral health, including substance use and abuse disorders, are often referred to as chronic diseases based on their pattern of service use over a lifespan (Centers for Disease Control and Prevention [CDC], 2016). The medical costs associated with caring for those with mental and behavioral health problems account for 86% of the nation's total health care costs, estimated as $2.9 trillion each year (U.S. Department of Health and Human Services [DHHS], 2017, p.159). In one of the most recent congressional budget requests proposed to the CDC by DHHS (2017), statistics such as these precede further written justification for a $50 million dollar increase to designate to mental health initiatives. The initiatives consist of plans to increase infrastructure for improving physical access to mental health care for children and adults (DHHS, 2017).

An increase in number of mental health clinics may have been requested because in the United States, one of the largest growing health trends is the number of emergency department (ED) visits for mental health related concerns (CDC, 2016; Simon & Schoendorf, 2014). Between 2006 and 2011 alone, the number of ED visits and hospitalizations for youth with injuries unrelated to mental health stabilized. In contrast, the rate for those who presented to the ED primarily with a mental health concern increased significantly by 21% (Torio, Encinosa, Berdahl, McCormick, & Simpson, 2015). The largest increases in ED visits were observed among boys aged 5 to 9 years

old with suicidal ideation (88% increase), and 1- to 4-year-olds with behavioral issues (61% increase). Inpatient hospitalizations due to suicidal behavior and self-injury amongst children increased by 104%, with the most dramatic spike (151%) observed amongst 10 to 14-year-olds (Torio et al., 2015).

Unsurprisingly, costs associated with mental health care have increased over time, approximately by 5 times the amount in millions spent in 1986 (CDC, 2016). For instance, from 2009 to 2014, estimates approximated a 22% increase (i.e., roughly 145 million to 186 million) in health care expenditures due to mental health treatment alone (operationalized as inpatient, outpatient care and prescription drug costs), excluding substance abuse treatment (CDC, 2016). By far, the largest increase was noted for outpatient care. This is not due to, for example, greater allocation of Medicaid dollars to outpatient clinics, but rather because greater and more varied services are being provided by ED facilities and intensive outpatient hospital programs, redesigned to lessen the burden traditionally shouldered by inpatient hospitals. Furthermore, in the review conducted by Torio and colleagues (2015), 90% of children who presented to the ED between 2006 and 2011 were labeled as "treat and release," meaning these families were sent home after a brief assessment found them not at risk to themselves or others. Children with a mental or behavioral health concerns were estimated to visit the ED 2.6 more times in one year, versus less than one more time (0.6) by the children who first presented without a mental health complaint (Torio et al., 2015). Medicaid absorbed the majority of these costs (CDC, 2016; Holder, Rogers, Peterson, & Ochonma, 2017; Torio et al., 2015), which today would approximate to $2,067 per child per visit, without adjusting for rise in health care costs separate from inflation.

Taken together, statistics may lead one to draw the following inferences. First, the large majority of children who present to the ED with mental or behavioral health problems exhibit a

level of severity that translates to a low level of clinical concern (i.e., does not indicate hospitalization or further action from a health professional). Second, these cases make up a subset of children who are likely to return the ED within the same year for similar reasons. These cases could be received and treated without the higher costs associated with accessing a hospital system, such as by an outpatient provider with specialized training. This way, families' problems are addressed and monitored, lowering the likelihood of a crisis event where parents present to tertiary levels of medical care. A more preventative approach to care such as this is not a novel solution, as reviews of the literature have long documented the savings in costs with outpatient care versus hospitalization for mental health specifically (Gabbard, Lazar, Hornberger, & Spiegel, 1997; Leslie & Rosenheck, 1999).

**STATE OF MENTAL HEALTH SERVICES FOR YOUTH.** The cost saving benefits of outpatient psychotherapy extend to children, as well, as its use has been found to lead to a lower rate of medical service utilization (Gabbard et al., 1997). Therefore, to many, the answer for reducing psychological suffering is to increase the amount of mental health service providers and outpatient clinics, under the assumption that greater availability will reduce need (DHHS, 2017; Holder et al., 2017). Others have requested that attention be paid first to our nation's current state of service provision and question whether increasing the quantity of providers will be sufficient (Bickman & Noser, 1999).

The most recent meta-analysis of psychological therapies for youth reviewed 447 randomized controlled trials (RCTs) that included 30,431 youth (Weisz et al., 2017). Results concluded that for children with anxiety ($d$ = .61) and conduct problems ($d$ = .46), efficacious treatments have been developed (Weisz et al., 2017), as well as specific practice elements that likely correspond to the mechanisms of change within these treatments (Bearman & Weisz, 2015;

Chorpita, Daleiden, & Wesiz, 2005). Unfortunately, the most robust practice elements of psychological treatments for youth have consistently been found to be absent from treatment-as-usual, or the type of services delivered in community settings (Borntrager, Chorpita, Higa-McMillan, Daleiden, & Starace, 2013). For example, studies of children treated in usual care for anxiety and disruptive behavior problems (the most common reasons for referral), have noted that evidence-based practice elements (e.g., exposure for anxiety) were absent in approximately 80% of individual therapy sessions (Borntrager et al., 2013; Garland et al., 2010).

Although the research on mental health treatment for youth has helped to develop disorder-specific treatment strategies that work, the deployment of these treatments to real world settings has been much less successful. Patients treated in efficacy studies also hardly match the characteristics of the children who are referred for services in community settings (Bearman & Weisz, 2015; De Haan, Boon, De Jong, Hoeve, & Vermeiren, 2013). These youths have more comorbid and severe symptom profiles and face various environmental barriers compared to children recruited in controlled trials (Bearman & Weisz, 2015; Borntrager et al., 2013; Costello et al., 2003). Differences may account in part for the low average rate of recovery documented throughout the literature (Bickman, 2008; Cooper et al., 2013; Smith & Jensen-Doss, 2017; Warren et al., 2010), and that 45-60% of children who enroll into outpatient services drop out before treatment completion (De Haan et al., 2013). Efforts toward increasing access to care are unlikely to help improve outcomes if the services provided are ineffective (Bickman, 2008), and unable to retain those who present for care (De Haan et al., 2013). In other words, more practitioners and greater availability of services may not make a discernable reduction in the psychological suffering of our children and families.

## Quality Assurance

**IMPROVEMENT OF SERVICES.** An alternative approach to addressing the high cost of low impact care is improving the quality of services delivered through existing outpatient settings. According to the National Committee for Quality Assurance, health care is improved through a) identification of standards, and b) development of performance measures. Once combined, comparisons between standards and performance allow accountability as areas for improving practice are illuminated. Quality assurance was first introduced for mental health care as the continuous quality improvement (CQI) approach (Bickman & Noser, 1999). CQI includes the assumption that a link exists between care processes and patient outcomes, and data collected from patients over time informs a clinician's psychotherapeutic practice. Once a clinician adopts a new skill, measuring its benefit through patient report is believed to support the clinician's permanent adoption of that skill for future practice (Bickman & Noser, 1999).

"Patient-focused" research emerged to evaluate the processes involved with CQI, where individualizing treatment to each patient is thought to improve mental health care (Lambert et al., 2002a; Sapyta, Riemer, & Bickman, 2005). This shift within the research field was a conscious one and rose from a concern that years spent promoting empirically supported treatments (ESTs) may have led to the misperception that mental health care services should be standardized in real world settings and perfectly match efficacy trial conditions (Ng & Weisz, 2016; Sapyta et al., 2005). The decision to begin to adhere to a patient-focused approach for research and practice did not signify a break from ESTs, but instead a collective response to work toward delineating a path for translating research findings to individual cases (Bearman & Weisz, 2015; Lambert, 2015; Weisz, Nelson, Burlingame, & Mondragon, 2012). One of the most promising ways to personalize

care includes the systemized use of progress monitoring and feedback, in conjunction with practice elements supported by the evidence base (Bearman & Weisz, 2015; Ng & Weisz, 2016).

**PERSONALIZATION OF CARE.** Over a decade ago, the APA Presidential Task Force (2006) published a report that defined and described the evidence-based practice (EBP) of psychology. Within the report, authors outlined the vantage of science-informed care and acknowledged that there will be many patient problems and clinical situations where the available evidence does not discern a specific approach. Psychologists must use clinical judgement and acquired expertise in such circumstances (APA, 2006). Clinical expertise is achieved through developing competencies that permit one to interpret the evidence and apply the best available solution to an ambiguous situation. Since this process will always involve probabilistic inferences, "ongoing monitoring of patient progress and adjustment of treatment as needed are essential to EBP of psychology" (APA, 2006, p. 276). In this report, clinician's use of a measurement feedback system to inform practice became an evidence-based practice element.

## Measurement and Feedback Systems

"Progress monitoring and feedback" is an intervention strategy that involves two processes: a) the repeated collection of patient-reported information relevant to treatment processes and/or psychological symptoms, and b) a composition of these data that are fed back to the clinician systematically and intermittently over an episode of care. "Progress monitoring and feedback" is a phrase synonymous with "feedback-informed treatment," "routine outcome monitoring (ROM)," and derivatives of these (e.g., "patient reported outcome measures" [PROMS]). It also represents the core intention of a measurement feedback system (MFS) (Edbrooke-Childs, Wolpert, & Deighton, 2016). Reference to an MFS adds specificity to the way in which data are collected and formatted into feedback for clinicians, through use of measurement tools that are, "administered

regularly throughout treatment to collect ongoing information about the process and progress of treatment, then provide[s] an automated presentation of the information… to mental health providers about their cases" (Lyon, Lewis, Boyd, Hendrix, & Liu, 2016, p. 442). Using an "MFS" is how progress monitoring and feedback is most commonly referred to in the United States, where emphasis is more often placed on the system that delivers the feedback. "ROM" and "PROMS" are used more frequently in the United Kingdom where measurement is the primary focus (Edbrooke-Childs et al., 2016).

For simplicity, "use of an MFS" will be used to refer to the entirety of literature that has examined the impact of progress monitoring and feedback as an intervention strategy in psychotherapeutic practice. Lyon and colleagues (2016) define an MFS as a "digital technology" where input and presentation of patient data is done through a commercialized platform that can be embedded into an electronic medical system or stand alone as an online and cloud-based entity (p. 442). It is important to note that the majority of studies that make up the progress monitoring and feedback literature have not used digital technology, especially older studies that employed less innovative and costly methods to monitor progress and deliver feedback. However, to be included in the current study's review, feedback must include data from past sessions and consolidated into graph form allowing simple and efficient analysis of patient progress. Some MFSs follow a graph with a set of recommendations intended to guide the next session (Lambert, Hansen, & Harmon, 2010), but this feature is not always included.

In summary, an MFS intervention includes both a systemized platform (digitized or non-digitized) and a patient-clinician process, for the purposes of the current study's review. Patients complete standardized questionnaires that are then consolidated into a graph that communicates patients' progress since treatment began, in real time, to the provider. The feedback platform

viewed by the clinician is often called a "dashboard" and may include additional features such as highlighted critical items endorsed by the patient (e.g., suicidal ideation), alerts or flags that communicate concern about the patient's status, and/or recommendations for next steps in treatment. Ultimately, the clinician decides whether and how often to access this information. See Figures 1-4 for exemplars of each of these components from one digitized MFS.

**MFS EFFICACY AND EFFECTIVENESS.** The act of progress monitoring and feedback has been stated to be an integral component of an evidence-based approach to therapy (APA, 2006), and is also listed as a "demonstrably effective" element of therapy relationships by the APA Task Force on Evidence-Based Therapy Relationships (Norcross & Wampold, 2011, p.127). Aside from these standards, progress monitoring and feedback represents an intervention strategy that does not belong to a specific theoretical orientation nor holds allegiance to any faction of therapeutic practice. Even without use of an EST, MFS use has demonstrated to improve treatment outcomes through "adapt[ing] psychotherapy to the *particulars* of the individual patient but do[ing] so according to *generalities* identified by research" (Norcross & Wampold, 2011, p. 131). The large majority of MFS studies that have been conducted include adult patients (Lambert & Shimokawa, 2011; Shimokawa, Lambert, & Smart, 2010; Lambert, Whipple, & Kleinstäuber, 2018). Within the last decade, studies conducted with youth as participants have emerged, heeding the recommendations set forth by the preceding adult research on study design and MFS technologies.

Mean improvement in outcome scores. Shimokawa and colleagues (2010) reviewed 6 RCTs that evaluated the effects of providing feedback to clinicians through repeated collection of the Outcome Questionnaire-45 on client outcomes (Lambert et al., 1996; Lambert et al., 2010). Participants were university college students who were randomized either to receive regular psychotherapy or psychotherapy with an MFS, and clinicians were doctoral trainees. Those in the

feedback condition experienced a significant reduction in distress and increased odds of improvement from baseline to the last treatment session compared to patients whose clinicians did not receive feedback ($d$ = .12) (Shimokawa et al., 2010). Knaup, Koesters, Schoefer, Becker, & Puschner (2009) reported similar findings in their review of 12 studies that both included RCTs and non-randomized designs. Feedback effects were reported to be larger when both clinician and patient received feedback on patient progress ($d$ = .30), compared to when only the clinician received feedback ($d$ = .09). Methodology concerning what constituted an MFS varied greatly across studies (Knaup et al., 2009), compared to the former review (Shimokawa et al., 2010), which only included studies that utilized the same patented MFS.

A more recent review that included 12 RCTs returned more tempered results for effects of an MFS on mean score improvement (Kendrick et al., 2016). The review included studies from both mental health and primary care settings, and three studies from Shimokawa and colleagues (2010) that met researchers' criteria for randomization (Kendrick et al., 2016). A significant effect of feedback on outcome scores was not found, with authors acknowledging the difficulty in making comparisons due to the heterogeneity across studies. Authors did find that when more severe patients were extracted from samples and analyzed separately, MFS use did significantly improve outcomes by the end of treatment ($d$ = .22) (Kendrick et al., 2016). Thus, the effects of an MFS were significant for the most severe patients (who also face the highest risk of premature drop-out) but less detectable for patients who began treatment with more mild concerns or respond to treatment as expected early on (Kendrick et al., 2016). This finding has been reported by others as well (Gondek, Edbrooke-Childs, Fink, Deighton, & Wolpert, 2016; Lambert et al., 2003).

Another cumulative review comprised of 32 studies used a broader and more inclusive definition of an MFS and reported similar findings (Gondek et al., 2016). More severe and

intractable patients were more likely to benefit from their clinicians' usage of progress monitoring and feedback when compared to less severe counterparts and all patients within a control condition, as found in 73% of their sample of studies (Gondek et al., 2016). Importantly, not only does severity at intake predict improvement in response to feedback (Janse, De Jong, Van Dijk, Hutschemaekers, & Verbraak, 2017; Smith & Jensen-Doss, 2017), but patients who exhibit an unexpected pattern of response to treatment are also more likely to benefit from their clinician receiving feedback (Gondek et al., 2016; Lambert et al., 2003; Shimokawa et al., 2010). Patients who either do not improve at a rate expected given their initial score at baseline or begin to deteriorate further from baseline are categorized as "not-on-track" (NOT) or "off-track," as they are at the highest risk for dropping out of treatment before their symptoms improve (Lambert et al., 2003; Shimokawa et al., 2010). Many MFSs include an embedded signal-alarm protocol that deploys electronic and color-coded notifications that alert clinicians to NOT patients (Delgadillo et al., 2017; Lambert et al., 2010; Lucock et al., 2015).

Treatment efficiency. For community care settings, treatment length and number of sessions is often a variable of interest due to the need to prioritize cost-effectiveness (Bickman & Noser, 1999; Brown & Jones, 2005; Kendrick et al., 2016). Many studies include average number of sessions as an outcome variable for an entire sample, but measurement of actual treatment *efficiency*, opposed to treatment *length*, has been thought to require different conceptualizations for NOT and "on-track" (OT) patients (Gondek et al., 2016; Lambert et al., 2003). NOT patients, due to their severity and unpredictability, are encouraged to remain in treatment until they improve. OT patients (i.e., those who respond to treatment as expected) typically require a less intensive program of care with fewer sessions to reach improvement status, compared to NOT patients

(Lambert et al., 2003). For the OT subgroup, more sessions do not predict a higher likelihood for improvement as they do for the NOT subgroup (Delgadillo et al., 2017).

Lambert and colleagues (2003) reviewed three seminal studies and concluded that feedback not only provided value to all patients who present for treatment, but that clinician use of an MFS makes treatment more efficient. OT patients in the feedback condition received approximately a single session less than the amount received by OT patients in the control condition. NOT patients in the feedback condition received 1 to 1 ½ more sessions than NOT patients in the control condition (Lambert et al., 2003). Authors concluded that feedback benefits NOT patients by keeping them in treatment longer in order to experience improvement, and OT patients benefit through shortened treatment episodes (Lambert et al., 2003). The differences due to condition led Lambert and colleagues (2003) to further infer that if clinics took advantage of an MFS for OT patients, clinicians would be able to terminate earlier with these patients, and shorter treatment episodes for this subgroup would compensate for the additional costs associated the NOT subgroup (not accounting for the costs absorbed by other care providers).

Lambert and colleagues' (2003) findings have been replicated (Gondek et al., 2016; Janse et al. 2017; Kendrick et al., 2016). Janse and colleagues (2017) conducted a longitudinal study with historical controls in a community setting where an MFS was implemented at midpoint with 1,006 patients and 85 therapists participating. Significantly more patients who enrolled into services after the MFS was implemented clinically improved within the first five sessions of treatment (Janse et al., 2017). The OT subgroup, particularly those diagnosed with depression versus any other disorder, also received two fewer sessions per treatment episode than did the matched controls who were served prior (Janse et al., 2017). Across the studies that report more sessions for NOT patients, they also report that all patients in a feedback condition benefit (OT

patients included) from use of an MFS compared to patients in a control condition (Gondek et al., 2016).

Delgadillo and colleagues (2017) examined the effects of an MFS+ condition versus an MFS already in use within a community care setting on treatment efficiency. In this study, the MFS was defined as the repeated collection of two short measures for depression and anxiety which were then electronically fed back to the clinician in graph form. The experimental condition (MFS+) included the addition of expectancy curves to signify OT status and alarm-signals to the clinician for when a patient became NOT (i.e., features often included in a basic MFS condition in other studies). Average cost of treatment was reported to be significantly higher for those in the control condition compared to patients in the MFS+ condition (Delgadillo et al., 2017). Authors inferred that higher cost related to the greater number of sessions for patients in the control condition, because clinicians did not have a reliable method for assessing progress and whether a patient should to be referred to upper level care (Delgadillo et al., 2017). Providing expectancy curves to clinicians to help detect individual treatment status may drive an MFS's impact on treatment efficiency. Although Delgadillo and colleagues' (2017) study does not provide a clear picture of treatment efficiency (a non-MFS control condition was not included), it does move the literature toward illuminating potential mechanisms for cost effectiveness.

Rate of recovery. In addition to symptom reduction and treatment length, the rate in which a patient reaches improvement status is relevant for judging the utility of an MFS. In one study conducted in mostly private practice offices, clinician use of an MFS did not affect symptoms overall, but it did associate with a faster rate of recovery (De Jong et al., 2014). In this case, feedback was delivered to both clinician and patient, not just to the clinician (De Jong et al., 2014). Similarly, in a community mental health care setting, Janse and colleagues (2017) found that

significantly more patients in the MFS condition achieved reliable positive change within the first five sessions of treatment, compared to those in the control condition (Janse et al., 2017). Additionally, one RCT conducted with university college students found that patients with clinicians assigned to the MFS condition reported larger gains in therapeutic alliance earlier in treatment than did patients in the control condition (McClintock et al., 2017). The use of an MFS may not only move a patient faster to recovery, it may also help foster the development of the therapeutic relationship.

Dropout prevention. The dropout rate for youth in community settings has been estimated to be 45%, according to clinician report (De Haan et al., 2013). Through use of an MFS, clinicians have the ability to predict whether an adult patient will stay in treatment and potentially later improve or deteriorate and preemptively drop out of treatment (Kendrick et al. 2016; Shimokawa et al., 2010). There is reason to believe these findings could extend to youth in treatment, as well. For instance, one sample of 2,310 youth were seen by 115 clinicians across one community mental health setting and one managed care setting (Warren, Nelson, Burlingame, & Mondragon, 2012). Approximately half of the sample ended treatment without significant improvement in symptoms, with fewer patients improving if treated in the community setting (Warren et al., 2012). If the same measurement tool used in the study also fed results back to each clinician in real time (instead of just being collected), 19% of the youth served in the community mental health care settings and 13% in the private practice settings that were deteriorating while in treatment would have been detected by their clinicians (Warren et al., 2012). With access to patient feedback, it is possible that dropout could have been prevented. Systemized feedback to clinicians has great potential to improve psychotherapeutic practice and retain those who are not progressing in care.

**MFS EFFICACY AND EFFECTIVENESS IN YOUTH POPULATIONS.** In order to assess whether the benefits of an MFS could transfer to youth, Bickman, Kelley, Breda, de Andrade, & Riemer (2011) conducted a multi-site RCT with 340 youth. Community mental health clinics that expressed interest in participating ($n$ = 28) were randomized to one of two conditions: an active condition where patient outcome data was fed back to clinicians, or a control condition where clinicians only received a summary of patient progress every 90 days (Bickman et al., 2011). Clinicians in the feedback condition received patient data within one to two weeks after collection and also received a summary of progress over the previous 90 days. Repeated measurement of youth symptoms and functioning was collected using clinician, child, and parent report at the end of each session. Data were collected via paper and pencil, which then researchers uploaded an electronic MFS, called the Contextualized Feedback System (Bickman et al., 2011).

Bickman and colleagues (2011) found that youth were more like to improve from baseline at a significantly faster rate if their clinician received more frequent feedback, according to youth ($d$ = .18), clinician ($d$ = .24), and parent report ($d$ = .27). According to intent-to-treat analyses, effects continued to remain significant according to youth and clinician report but not for parent report. Interestingly, a significant dose-response relationship between number of feedback reports viewed by the clinician and outcomes was found (Bickman et al., 2011). The more a clinician viewed the patient's feedback report, the greater the gains the patient made in treatment (Bickman et al., 2011). Once clinicians who never accessed the MFS (despite being assigned to the feedback group) were removed from analyses, effect sizes increased for MFS's impact on rate of recovery according to clinician report ($d$ = .40) and youth report ($d$ = .27) (Bickman et al., 2011).

Bickman and colleagues (2016) conducted another study to evaluate the effects of an updated version of their original MFS (Bickman et al., 2011). Youth ($n$ = 257) were randomized

to the MFS or control condition within two outpatient clinics ("Clinic R" and "Clinic U"), where clinicians ($n$ = 21) saw youth in both study conditions (i.e., randomization at the patient level). Clinicians could immediately access feedback electronically for patients in the MFS condition, versus every six months for their patients in the control condition. Both parents and youth reported a significant reduction in symptoms over the course of treatment, regardless of condition and site. The feedback effect was significant according to clinician report at Clinic R but not at Clinic U (Bickman et al. 2016). At Clinic R, the dose-response relationship among clinician access to feedback and symptom improvement (Bickman et al., 2011) was found to be significant for to clinician and youth report (Bickman et al., 2016), i.e., the more clinicians viewed the feedback, the greater an effect it had on improvement. Implementation was found to vary greatly across the two sites, which opened questions related to treatment integrity (Gleacher et al., 2016). Differences across sites included administration procedures, stakeholder buy-in, and level and type of ongoing consultation and supervision (Gleacher et al., 2016). This highlighted the importance of MFS integrity, specifically adherence, discussed next.

**ADHERENCE TO THE MFS.** In an attempt to quantify the differences across the two sites, Bickman and colleagues (2016) developed a composite variable to reflect the amount of intervention dosage received by the clinician (i.e., dose-response relationship), and called it the Implementation Index. The index equaled the number of times a clinician accessed feedback via the electronic MFS divided by the total questionnaires completed for a given case (Bickman et al., 2016). Although the index does not quantify how feedback was used by the clinician, it provides a unit related to MFS engagement that can be used to compare clinicians to each other by calculating the degree to which the MFS was accessed. On average, the indices for clinicians at Clinic R averaged 25% higher than indices of clinicians at Clinic U (Bickman et al., 2016). More

interestingly, analyses showed that if a clinician adhered to the intervention and accessed the MFS at a rate of 75% or above, the effect sizes would have increased to .70 (according to youth) and .80 (according to clinician), respectively, only after 10 weeks of treatment (Bickman et al., 2016). Thus, the more clinicians access the MFS, the more patients benefit. Clearly, clinician behavior greatly affected outcomes, as seen in past studies with adult patients (Amble et al., 2015; De Jong, van Sluis, Nugter, Heiser, & Spinhoven, 2012; Lutz et al., 2015).

Bickman and colleagues (2011) provided some support that children and adolescents benefit when an MFS is combined with standard treatment-as-usual delivered by community providers. Both studies highlighted that the presence alone of an MFS or merely the repeated collection of data is not enough; clinicians *must* access the feedback in order for an MFS to add value to practice (Bickman et al., 2011; Bickman et al., 2016; Kendrick et al., 2016). Approximately 35% of clinicians assigned to the feedback condition in Bickman and colleagues' (2011) study did not access the feedback reports. The clinicians who did were found to discuss relevant symptoms in following sessions with patients more often (Kelley, de Andrade, Bickman, & Robin, 2012) and addressed specific problem areas faster than clinicians in the control condition did (Douglas et al., 2015). For example, an adolescent who endorsed recent substance use led to the MFS issuing a warning signal to the clinician. This topic was then more likely to be brought up and discussed with the adolescent, according to the clinician, during a following session (Douglas et al., 2015). When a crisis of any kind was reported and fed back to the clinician, the treatment processes and symptoms related to functioning were more likely to be discussed in a following session (Kelley et al., 2012). Whether it was the clinician or the patient who broached the topic in session cannot be determined, but session content individualized to patient was not reported by clinicians without access to feedback (Douglas et al., 2015; Kelley et al., 2012).

## The State of the Evidence for MFSs

The progress monitoring and feedback literature for mental health practice has progressed greatly over the past ten years with advancements in data collection innovation and feedback efficiency. But unlike an EST that is first developed, manualized, and then tested through efficacy trials, progress monitoring and feedback as an intervention strategy does not have a list of criteria that establishes it as a distinct entity from other intervention strategies. Even though standards have been promoted (Lyon et al., 2016), any kind of monitoring and feedback procedure is permitted to be called and tested as an "MFS." Because of this, interventions within the studies that make up this literature vary greatly (Bergman et al., 2018; Kendrick et al., 2016). Without agreement, it is challenging to compare different MFS systems and confidently interpret the findings (Carlier et al., 2012; Gondek et al., 2016; Knaup et al., 2009). The most significant differences lie in feedback conceptualization and data collection method.

**FEEDBACK DEFINITIONS.** For the purposes of the current study proposal, feedback has been conceptualized as a compilation of information that has been previously collected from a patient with the idea that the assigned clinician will find the information useful for planning future sessions (Bearman & Weisz, 2015; Sapyta et al., 2005). This assumption is not unanimous among the MFS literature, for instance, when feedback measures are used as an outcome variable in group therapy. Crits-Cristoph and colleagues (2010) examined the impact of a semi-automated MFS on alliance, substance use, and satisfaction with group therapy treatment and reported no differences across experimental and control conditions. Similar findings were reported in Schectman and Sarig's (2016) study that evaluated an MFS with group therapy in the school setting. In both studies, feedback was aggregated across reporters and delivered to group leaders as an average; therefore, feedback was conceptualized as a marker of group process instead of an individual's progress

through treatment. Additionally, though feedback was treated as a repeating variable over time, many patients and students provided feedback only once (Crits-Cristoph et al., 2010; Schectman & Sarig, 2016), meaning change over time was not connected to an individual but to the collective group environment. This type of feedback appears to differ from other studies mentioned up to this point, yet these studies and similar ones are included in cumulative reviews.

The form in which feedback is compiled and presented to the clinician is likely to affect understanding and application of the information. Since the criteria for what equates feedback has not been defined, disagreement on which components are integral is reflected across study design (Kendrick et al., 2016). For instance, simply a graph of an individual patient's progress across time has represented both the feedback (Reese, Norsworthy, & Rowlands, 2009a; Reese et al., 2009b; Stein, Kogan, Hutchison, Magee, & Sorbero, 2010) and control condition (Delgadillo et al., 2017, 2018), as well neither conditions (Janse et al., 2017) . Feedback in some studies has been summarized through interactive and computerized platforms, a mode with most empirical support for eliciting action (Harkin et al., 2016; Kluger & DeNisi, 1996). These platforms have evolved to store data online, graph patient progress, and provide treatment planning guides or next step recommendations based on a patient's trajectory (Boswell et al., 2015; Lambert et al., 2010).

Some MFSs employ a varied notification system where green, yellow, and red signify improvement, no significant change, and deterioration, respectively, presented alongside graphs electronically (Lambert et al., 2010), through a traffic light design (Lucock et al., 2015), or a colored sticker placed onto physical patient files by researchers (Lambert et al., 2002a, 2002b). Connolly-Gibbons and colleagues (2015) collapsed patient categories to green and red (i.e., progressing as expected or not) and issued printouts to clinicians before every session. This type of color coding and additional notifications to clinicians for when patients go "off track" are

thought to be what makes an MFS effective for patients (Delgadillo et al., 2017; Douglas et al., 2015; Lambert et al., 2002b), but further study into the feature(s) that best facilitates feedback's utility for clinicians is warranted.

**PROGRESS MONITORING PROCEDURES.** Without specification on how often feedback should be delivered, data collection procedures vary across studies that claim to be measuring the same intervention approach. Patient data should be collected often and fed back immediately before the patient or clinician has terminated treatment (Bickman & Noser, 1999; Lyon et al., 2016; Sapyta et al., 2005), but this may not have been the case in some studies. Some studies have included a system that delivered feedback at on a fixed timetable (e.g., once a month) that was completely separate and unrelated to appointment schedules and data collection frequency (Brown & Jones, 2005; Davidson et al., 2017). Such a design proved to be insufficient, as many patients deteriorated and left treatment before clinicians received the feedback (Brown & Jones, 2005). Other feedback conditions have included designs where patients reported their progress every 5 (De Jong et al., 2012), 10 (Warren et al., 2012), or 20 sessions (Lutz et al., 2015).

One of the largest differences across studies involves the *way* in which data is collected, and idiosyncrasies concerning data collection have the potential to bias patient response. Indeed, one study found that an independent measure collected in addition to the progress measures administered through an MFS returned different results concerning intervention effectiveness (Janse et al., 2017). Janse and colleagues (2017) found that use of an MFS led to significantly improved symptoms according to the repeated measures that were administered to patients ($n$ = 1006) in front of and then immediately scored by clinicians ($n$ = 85) at the beginning and end of each treatment session. Similar findings were not reflected by the secondary measure collected at pre- and post-study by researchers. Authors mention that the measures scored in front of patients

were positively skewed, i.e., the highest choice option was endorsed in succession and on average after the third session (Janse et al., 2017). Similar results have been found in other studies that used a similar MFS procedure (Reese et al., 2009a, 2009b). It is possible that other processes are being measured in this scenario, aside from the intervention's impact on a clinician's practice or patient's functioning (Bergman et al., 2018; Kendrick et al., 2016).

A clinician's role in the data collection process differs greatly across studies. How data are collected is a crucial feature to mention, as procedures surrounding use of an MFS greatly affect clinician buy-in and its adoption when implemented with community clinicians (Gleacher et al., 2016). Within one study, the feedback condition included one site where data was collected via paper and pencil and another site that permitted patients to input their data electronically (Bickman et al., 2016), which is often preferred because of efficiency and lower chance of data entry error (Hall et al., 2014; Hatfield et al., 2010; Lyon et al., 2016). Clinician responsibilities required for each MFS vary greatly and include scoring measures by hand and then immediately discussing results (Miller, Duncan, Sorrell, & Brown, 2005), faxing patient responses to researchers in real time (Stein et al., 2010), inputting the data themselves once the session has ended (Hatfield et al., 2010), or delivering items out loud (Kendrick et al., 2016). When an MFS's functionality is dependent upon extra work from clinicians, implementation suffers (Gleacher et al., 2016; Kotte et al., 2016). Though, study designs that force clinicians to attend to patient scores in any capacity as part of the data collection process are likely to influence intervention dosage. The intervention dosage is likely to be greater for clinicians involved in data collection compared to clinicians who must go out of their way to access data when assigned to a feedback condition (Harkin et al., 2016).

In addition to frequent monitoring and timely feedback, the actual measures that are administered through an MFS deserve equal weight of evaluation (Bickman & Noser, 1999;

Harkin et al., 2016). Clinicians favor measures that are idiosyncratic to the patient (Bickman et al., 2000; Jensen-Doss et al., 2017), yet broad enough to be relevant at each time point that they are administered (Hall et al., 2014; Sapyta et al., 2005). The majority of the measures administered through MFSs have been tested and standardized in prior studies, but some studies included in meta-analytic reviews chose to design and implement a new measurement tool for their specific site (Connolly-Gibbons et al., 2015; Stein et al., 2010) before psychometrics were established. Lastly, measures should be brief but sufficiently comprehensive to be sensitive to change over time (Harkin et al., 2016; Kelley & Bickman, 2009). At this time, the three most popular MFSs repeatedly administer measures that contain 8 to 64 items at each time point (Barkham et al., 2001; Miller et al., 2005; Lambert et al., 1996).

## Theoretical Explanations of Progress Monitoring and Feedback

Reviews of studies conclude that patients in a feedback condition experience greater symptom reduction than patients in a no-feedback condition (Knaup et al., 2009; Lambert et al., 2003; Shimokawa et al., 2010). Effects of an MFS have been found to be particularly beneficial for patients with the highest risk for dropout from treatment (Gondek et al., 2016; Kendrick et al., 2016). Specific characteristics that make an MFS more or less effective for clinicians and their patients have not been defined; thus, the active ingredient(s) of progress monitoring and feedback is still largely unclear. Proposed hypotheses vary widely, borrowed from adjacent philosophical fields and adapted for human psychology. The most widely discussed theory used to help explain why an MFS may be effective is the Feedback Intervention Theory (Kluger & De Nisi, 1996). In addition, it has been proposed that an MFS's impact is mediated by the therapeutic alliance between clinician and patient (Gondek et al., 2016).

**FEEDBACK INTERVENTION THEORY.** Developed by Kluger & DeNisi (1996), the Feedback Intervention Theory (FIT) describes a series of decisions an individual makes, consciously or subconsciously, in response to feedback. A "feedback intervention," is defined as "actions taken by an external agent(s) to provide information regarding some aspect(s) of one's task performance" (Kluger & DeNisi, 1996, p. 255). The basic tenets of FIT state that, a) behavior is based on making comparisons of feedback to one's goals or standards, and b) an individual's attention is limited, so the largest gaps are attended to for priority. When feedback is presented, an individual makes the choice either to adjust one's behavior (or standard), or to reject the information (or reframe it). This decision is influenced by an individual's capacity at each of the three levels of control: task learning level, task motivation level, and meta-task level (Kluger & DeNisi, 1996). The task learning level of control refers to the details and steps necessary to acquire a novel and specific skill. Task motivation represents the middle level that receives all incoming information and decides whether to send incoming information to either of the two other levels. The meta-task level represents thinking about the self within the context of new information. See Figure 5 for an illustration.

For example, once a clinician receives feedback delivered via an MFS, this information lands at the middle level, the task motivation level. If this information is expected or negative in nature, e.g. "therapeutic alliance is poor," depending on the capacity at the middle level, the data is sent upward or downward (Kluger & DeNisi, 1996). If cognitive demands from elsewhere have usurped the sufficient motivation necessary to attend fully to the information, it is likely to be sent upwards to the meta-tasks level. There, framing effects based on an individual's innate personality characteristics and past experiences determine whether the feedback is rejected and reframed to be more acceptable to the individual, e.g. "This information is wrong because of the way in which

the questions are asked; My patient did not understand the question," (Kluger & DeNisi, 1996). The information could have also been sent downward to the task learning level, where upon further inspection the clinician is likely to accept the feedback and assess the steps necessary to address the issue, i.e., repair the therapeutic alliance. Theoretically, it is up to the supervisor to ensure that the conceptualization of the feedback, or the task details, as well as the steps to for addressing an issue are accurate (Sapyta et al., 2005; Tracey et al., 2014).

The main ideas promoted by FIT is that feedback interventions bring attention to important information that would normally go unnoticed, and under ideal circumstances, feedback can facilitate clinician learning and skill development (Kluger & DeNisi, 1996; MacDonald & Mellor-Clark, 2015; Sapyta et al., 2005). Goldberg and colleagues (2016) sought to test whether implementation of a novel MFS in a community mental health setting improved the clinician practice over time, versus only tailoring therapy for each patient individually. Patient outcomes were collected across seven years from 5,128 patients and 153 clinicians (Goldberg et al., 2016). In order to account for turnover and the possibility that better clinicians were hired as the study proceeded, employment start date was included as a statistical predictor of patient outcomes. According to outcome scores collected pre-post, patients were found to improve at a significantly higher rate at the seven-year mark from when the MFS was implemented ($d$ = 1.17), at a rate of 0.035 of an effect size each year, and clinician start date was not a significant predictor (Goldberg et al., 2016). Also, patient-reported symptoms at baseline increased each year, to which Goldberg and colleagues (2016) inferred that although patient severity on average increased over time, clinicians were becoming more skilled and effective in treatment delivery. This study is the first to date that lends some support to the proposed theory of FIT concerning the mechanism of an MFS.

**THERAPEUTIC ALLIANCE.** Norcross and Wampold (2011) introduced the "evidence-based psychotherapy relationship" in conjunction with meta-analytic evidence of the therapeutic alliance's impact on psychotherapy outcomes (Horvath, Del Re, Flückiger, & Symonds, 2011). Within their special issue on this topic reporting the results of the APA Task Force on Evidence-Based Therapy Relationships, common correlates of the therapeutic alliance and effective methods of adapting therapy were described. Along with foundational elements such as empathy, goal consensus, collaboration, and positive regard, collecting patient feedback through the use of standardized measurement tools was included as related to the development of the therapeutic alliance (Norcross & Wampold, 2011). Progress monitoring and feedback has been theorized to impact engagement and motivation for treatment (Carlier et al., 2012; Macdonald & Mellor-Clark, 2015), and through strengthening the components that make up the therapeutic relationship, the use of an MFS has been theorized to prevent premature termination and lead to better patient outcomes (Horvath et al., 2011; Lucock et al., 2015; Norcross & Wampold, 2011).

Unsurprisingly, one sample of 482 clinicians listed status of the therapeutic alliance as one of the most useful pieces of information to know when working with an adolescent patient (Bickman et al., 2000). More recently, researchers have been interested in how alliance may interact with a clinician's use of an MFS. Studies that included measures for both alliance and outcomes have largely reported that symptom improvement from the use of an MFS does *not* significantly relate to (nor hinder) the therapeutic alliance (Gondek et al., 2014; Grossl, Reese, Norsworthy, & Hopkins, 2014; Lucock et al., 2015; Reese et al., 2009b; Schechtman & Sarig, 2016). In one study that reported null findings concerning the effects of an MFS on treatment outcomes, use of the MFS was found to affect the rate of change in patient report of the therapeutic alliance (McClintock et al., 2017). The clinician and patient's relationship developed faster if the

clinician was assigned to the MFS condition (McClintock et al., 2017), so some mechanism associated with the process of monitoring patient progress impacted the therapeutic alliance in a positive way. In summary, if a strong therapeutic alliance is a chief concern to which the clinician believes symptom improvement is directly tied, then the use of an MFS has only been reported to help facilitate its development. Likewise, MFS use also improves treatment outcomes, although through a seemingly separate path.

## Statement of the Problem and Purpose

Medical and professional care devoted to mental health treatment has become more costly over time, largely due to the increase of youth accessing the most expensive levels of care (e.g., EDs) that may be ill equipped to treat mental and behavioral problems past the acute phase (Holder et al., 2017; Torio et al., 2015). This trend could be explained both by poor access to outpatient mental health care (e.g., waiting lists for intake appointments), as well as ineffectiveness of these services once enrolled (Bickman, 2008). Unfortunately, none of these services at any level are meeting the need in proportion to its growth (Costello et al., 2014; De Haan et al., 2013). Instead of efforts that target quantity of outpatient mental health treatment providers, quality of these entities must first be addressed (Bickman & Noser, 1999; De Haan et al., 2013). The research to date has championed practices that emphasize the use of progress monitoring and feedback in practice (Bearman & Weisz, 2015; Ng & Weisz, 2016). Although the specific mechanisms that facilitate an MFS's effectiveness are unknown, their use has been empirically proven to improve treatment outcomes, especially for intractable patients who traditionally leave treatment before recovery (Kendrick et al., 2016). Progress monitoring and feedback as an intervention strategy has been studied thoroughly in the adult literature while only a handful of studies report an MFS's impact on services for youth (Bergman et al., 2018). The purpose of the current study is to

contribute to this section of psychological science that aims to improve our existing mental health care systems for youth.

## Research Questions and Hypotheses

**RESEARCH QUESTION 1.** Does a clinician's use of an MFS impact treatment engagement for youth treated at a community mental health center?

***Hypothesis 1a.*** The use of an MFS will decrease dropout rate, and this relationship will be moderated by a patient's urgency at intake.

*Rationale 1a.* Since clinicians did not have access to an MFS that monitored patient progress, patient symptom deterioration was quantitatively undetectable. It is possible that because clinicians are responsible for an overwhelming number of tasks for each patient, they may be prone to error when predicting individual patient treatment response (Hannan et al., 2005; Hatfield et al., 2010; MacDonald & Mellor-Clark, 2015), and efforts to retain the patient may not have been made during Wave A. Urgency at intake is hypothesized to help explain the relationship between MFS use and dropout, since patients whose symptoms or presenting problem(s) were more severe at intake (versus less severe) may be more likely to drop out (De Haan et al., 2013), since pattern of response during treatment was not accessible in Wave A.

***Hypothesis 1b.*** The use of an MFS will increase treatment length per treatment episode, and this relationship will be moderated by urgency at intake.

*Rationale 1b.* Patients who begin at a more severe state at intake are more likely to improve through a clinician's use of an MFS versus patients who begin treatment at a lower level of severity, and this may be because the MFS keeps more severe patients in treatment longer (De Jong et al., 2014; Gondek et al., 2016; Kendrick et al., 2016; Janse et al., 2017). Access to the MFS may allow clinicians to make more informed decisions concerning termination of services, and

those who are more severe at baseline are more likely to be retained in treatment longer in order to reach improvement status (Brown & Jones, 2005; Cooper et al., 2013; Lambert et al., 2003).

**RESEARCH QUESTION 2.** For patients whose clinicians have access to an MFS, does their rate of symptom change depend upon intervention dosage?

***Hypothesis 2.*** Patients whose clinicians access their feedback more often will improve at a faster rate than patients whose clinicians access feedback less often, and this relationship will be moderated by urgency at intake.

*Rationale 2*. A dose-response relationship among number of times a clinician views patient feedback and patient outcomes has been found in past studies (Bickman et al., 2011, 2016). Clinicians who choose to view the feedback reports delivered through the MFS may be more likely to deliver more personalized care. Since a control group does not exist in this study, patients are expected to improve over time independent of the MFS intervention, but higher MFS dosage is expected to associate significantly with greater treatment outcomes (Bickman et al., 2011, 2016; De Jong et al., 2012). Furthermore, this effect is expected to be strengthened for those who presented for treatment at a higher symptom severity compared to patients who are less severe at baseline. Since MFSs are believed to be the most effective for NOT patients (Kendrick et al., 2016), urgency at intake is believed to moderate the relationship between intervention dosage and treatment outcomes, according to youth- and parent- report. Those with higher levels of urgency for treatment at intake are believed to have greater symptom severity and to be less likely to respond in a predictable manner to treatment.

Chapter 3

# METHOD

## Participants

Patient data for the current study were drawn from an electronic medical record database managed by one community mental health clinic (CMHC) for youth in Central Texas. The study's sample ($n$ = 538) was collected from an archival database of families ($n$ = 1,064) that presented for outpatient individual therapy between February 2016 and February 2018. Inclusion criteria included youth with at least one legal guardian(s), hereafter referred to as "parent(s)," who attended at least one individual therapy session after an intake appointment conducted on-site. The child or adolescent had to have attended the individual session either alone or with a parent present. The child or adolescent had to have been between the ages of 4 to 17 years old at time of intake. The youth and/or parents could be English or Spanish-speaking. Parents must have also demonstrated a fourth-grade reading level in order to be included into the study.

**Patient characteristics.** The study included two waves of patients: Wave A and Wave B. Participants from Wave A were not included in Wave B. Wave B patients must have ended treatment before the end of Wave B (February 2018). "Treatment episodes" is used to refer to a patient's treatment duration where individual therapy was conducted by the same clinician on an intended routine schedule (e.g., once a week, once every two weeks). A treatment episode has ended when the patient begins a transfer procedure to a new clinician, no-shows at least twice in a row, or attends a termination session. Youth who presented for an intake appointment only, group therapy, walk-in clinic, or home-based services delivered by clinicians employed by the CMHC were excluded from the study.

The final sample included 538 families who were served through individual therapy services across Wave A (*n* = 279) and Wave B (*n* = 259). Mean age at intake was 9.88 (*SD* = 3.90), 57% male (*n* = 307). The majority of patients (51%) identified as either Latinx or Hispanic (*n* = 277), 39% as White/Non-Hispanic (*n* = 208), 8% as Black/African American (*n* = 42), and approximately 1-2% as Asian American/Pacific-Islander or "Other" (*n* = 8). The majority of patients (84%) received care through some type of subsidized program, such as a federal or private grant, or government-provided insurance. See Table 2 more information on patient characteristics divided by wave membership.

**Clinician characteristics.** A second set of participants included mental health clinicians employed by the CMHC between February 2016 and February 2018 (*n* = 49). Of these 49 clinicians, 21 delivered services to youth in both Wave A and B. Twenty-eight clinicians delivered services only during Wave A or Wave B. This group is made up of mostly White/Non-Hispanic women (72%). Of those who were licensed (*n* = 37), the majority (81%) were masters-level clinicians. Clinicians were mostly social workers (n = 20), and the rest were trained either through doctoral school and clinical psychology programs (n = 18), or a counseling masters program (n = 11). Caseload averaged across wave equaled a median of 6 patients per clinician (*M* = 7.70, *SD* = 5.75). See Table 3 for additional clinician characteristics and separation of the above estimates separated according to wave.

## Procedures

**STUDY TIMELINE.** The study includes two waves: Wave A and Wave B. Wave A represents the year between February 15th, 2016 to February 15th, 2017; Wave B represents the year between February 16th, 2017 to February 16th, 2018. The midpoint between the two waves is when the measurement feedback system (MFS) was rolled out agency-wide within the participating

community center and satellite school-based clinics. All clinicians were mandated to begin to use the MFS and administer MFS measures for all patients at this point in time, February 15th, 2017. All clinicians employed by the CMHC at that point in time attended a one-day training delivered by the MFS developers three months prior to the implementation date.

All participants and their legal guardian(s) underwent notice of privacy and informed consent agreement prior to collection of protected health information and questionnaire data as part of standard clinic procedure. Clinic documents included notification that patient information may be used for quality assurance purposes. For the current study, data collection began in February 2018. Patient demographic data for those served between February 2016 to February 2018 (i.e. episode of care ended prior to the implementation of the MFS) were extracted from Filemaker, the agency's electronic medical record system. Questionnaire data for patients who were served only during Wave B (i.e. episode of care began following the implementation of the MFS) were extracted directly from the MFS's cloud-based platform, the OQ® Analyst (Lambert et al., 2010). The current study was classified as a program evaluation by University of Texas's Institutional Review Board and characterized as exempt for all data collection and analyses procedures (Study No. 2018-01-0164).

For patients in Wave B, measures associated with symptom severity were completed electronically by families once they arrived to the clinic immediately before meeting with their designated clinician. Both parents and youth aged 12 years or older completed a self-report measure on iPads provided by front desk staff. Each measure took approximately 3 to 5 minutes to complete. Since all data were collected electronically, clinicians were instructed to examine questionnaire responses via the online MFS platform immediately before beginning a therapy session. Clinicians were also trained by the MFS developers to use patient feedback to inform

clinical decision-making regarding whether a patient was progressing, or if a rupture in the therapeutic alliance was reported either by parent or youth.

**MEASUREMENT FEEDBACK SYSTEM.** The OQ® Analyst is the most widely studied online interactive platform that tracks patient response during ongoing mental health treatment (Lambert et al., 2010, 2018). The OQ® Analyst is a system that administers measures and then presents feedback to stakeholders, suitable for clinical use regardless of a clinician's theoretical orientation or plan for individual therapy (Lambert et al., 2010). Feedback reports are produced instantaneously after a parent or youth completes a questionnaire, allowing them to be reviewed by a clinician before an individual or joint session with both parent and child begins.

The OQ® Analyst provides color-coded alerts in white, green, yellow, and red to classify patient status at every point of measurement (Lambert et al., 2010). The system compares a given patient's rate of improvement with normative, expected rates of improvements based upon samples of patients at the same initial level of distress. A "white" alert notifies the clinician that the patient is functioning in the subclinical range. "Green" alert status lets the clinician now the patient is improving as expected and change in treatment plan is unnecessary at that time. "Yellow" alert status notifies the clinician that the patient is not making expected progress, and an alternate treatment plan should be considered. A "red" notification means definite insufficient progress and chances are high that the patient will drop out of therapy before treatment ends. Patients whose scores generate a red or yellow alert during treatment, the case becomes classified as "NOT" ("not-on-track"). Patient status can change from session to session, but information that informs status is accumulated from past measurement scores. See Figure 2.

## Measures

**DEMOGRAPHIC INFORMATION.** Demographic information for the parent and youth were collected at the intake appointment through a semi-structured interview measure constructed by the mental health agency. The following youth-specific variables were collected: age, gender, ethnicity, trauma background (i.e., "Were you ever a victim to violence?" Y/N), and financial options (grant eligible, insurance coverage, sliding scale, pay in full). One item is scored by the intake clinician on a 6-point Likert scale and refers to symptom severity at intake, or level of urgency based on presenting symptoms (1 = "low concern/severity," 6 = "very high concern/extreme severity").

Youth presenting problems and major symptoms were classified at intake or after the first therapy session through DSM-V diagnoses, which were later included in analyses as "internalizing" or "externalizing" symptom profiles. Children with an "internalizing" profile included those diagnosed primarily with any type of anxiety or mood disorder, including adjustment disorders with anxiety or depressed mood. Post-traumatic stress disorders and somatic disorders were included in this category. The "externalizing" category included youth with one or more of the following diagnoses, attention-deficit/hyperactivity disorder, oppositional defiant disorder or conduct disorder, adjustment disorder with disturbance of conduct, and any "other" or "unspecified" disorder characterized by disruptive behavior. If a youth presented with a mixed externalizing or internalizing profile, diagnostic type was categorized based on the primary diagnosis listed. The majority of patients (58%) exhibited a comorbid diagnostic profile. See Table 2 for patient data delineated by wave.

**TREATMENT ENGAGEMENT.** Classification of dropout and treatment length per treatment episode to reflect level of treatment engagement were collected for all patients who enrolled into

services and attended an intake appointment and then at least one treatment session with their assigned clinician. “Dropout” was defined according to recommendations set forth by De Haan and colleagues (2013). Patient were classified as having dropped out if they did not return to treatment after two “no shows” (i.e., no call, no show to scheduled appointment), or if they expressed desire to end treatment and refused a termination session with the clinician. If the patient or family member expressed desire to end treatment and attended a termination session, the patient was not classified as having dropped out. According to agency rules, if a family wished to return to treatment after a period of inactivity, the patient would have to attend another intake appointment, signaling the beginning of another treatment episode often with a new clinician. Deliberate transfer to another clinician, as in the case of a clinician leaving, counted as new treatment episode for one patient.

Treatment length was measured two ways, as number of sessions and days enrolled into treatment. Sessions were counted if attended by either the patient or parent and did not include the intake appointment conducted by an intake specialist. Whether sessions were held as scheduled or missed including the reason why (e.g., clinician cancelled within 48 hours) was collected from clinician treatment progress notes and billing records. **SESSIONS MOST OFTEN WERE SCHEDULED AT LEAST 24 HOURS IN ADVANCE BUT IMPROMPTU MEETINGS THAT LASTED 45 MINUTES OR MORE WERE ALSO COUNTED AS A SESSION. NEITHER PHONE CONSULTATIONS NOR MEETINGS WITH TEACHERS WERE INCLUDED IN SESSION COUNT. IN ORDER TO BETTER CAPTURE TRUE LENGTH OF A TREATMENT EPISODE, THE TREATMENT LENGTH VARIABLE WAS ALSO MEASURED THROUGH NUMBER OF DAYS PER TREATMENT EPISODE. IF MANY WEEKS PASSED BETWEEN SCHEDULED APPOINTMENTS, FOR INSTANCE, DUE TO HOLIDAYS OR CLINICIAN CANCELLATIONS,**

**SESSION NUMBER WOULD NOT REFLECT THE AMOUNT OF TIME EACH PATIENT REMAINED ON A CLINICIAN'S CASELOAD AND NUMBER OF DAYS IN TREATMENT WOULD.**

**SYMPTOM SEVERITY.** Two versions of the Youth Outcome Questionnaire-30 (Y-OQ) administered through the OQ® Analyst were used for parent and youth self-report of symptom severity or level of psychological distress (Burlingame et al., 2004). The parent version is indicated for report of children and adolescents aged 4-17, and the self-report version can be completed for those aged 12-18. The Y-OQ is a shortened version of the full Youth Outcome Questionnaire (Y-OQ 2.01; Burlingame et al., 2001), comprised of 30 items rated on a 5-point Likert scale from 0-4 (0 = "almost never," 4 = "almost always"). Summative scoring is used to produce a total overall score across six areas of emotional and behavior distress (somatic, social isolation, conduct problems, aggression, hyperactivity/distractibility, and depression/anxiety). Scores range from 0 to 120, with higher scores indicating greater psychological distress. Scores at or above the established clinical cutoff score of 29 are considered in the clinical range for level of distress (Burlingame et al., 2004). The Y-OQ has demonstrated sufficient inter-rater reliability, $\alpha$ = .92-.93 (Burlingame et al., 2004; Shechtman & Sarig, 2016), and sensitivity to change (.70-.80; Mosier et al., 2001). An additional five items related to therapeutic alliance were included at the end of the Y-OQ for both parent and youth to complete. The items have not undergone psychometric examination separate from the Y-OQ 30, and the derived data were not examined in the current study. See Appendices A and B for both reporter versions of the measure in paper form.

Both measures were administered to parents and youth of at least 12 years of age before each individual therapy appointment. Both reports were included in separate analyses since the best way, or method, for measuring whether a child is experiencing full criteria for a given disorder is through including perspectives from multiple reporters (De Los Reyes & Kazdin, 2005).

Agreement between parent and child is often low, especially for patients who are experiencing more internalizing systems related to anxiety and depression (Cooper et al., 2013; Hawley & Weisz, 2002). Disagreements among parent and child are often related to how observable a symptom is, e.g. aggressive acts are more observable than feelings of despair, and how distressing or problematic a symptom is perceived to be for the child, e.g. yelling or hitting is less of a problem for the child than the parent who is on the receiving end (De Los Reyes & Kazdin, 2005; Hawley & Weisz, 2002). Different perspectives, as well as the difference between them, provides unique information concerning a patient's trajectory in treatment.

**INTERVENTION DOSAGE.** To measure whether change in Y-OQ scores is influenced by the amount of feedback the clinician received, or intervention dose, an "Implementation Index" was calculated (Bickman et al., 2016). The composite considers two dimensions: the amount of Y-OQs completed per case, and how often the clinician viewed the feedback reports for that specific case. The OQ® Analyst can indicate whether the feedback has been accessed by the clinician but not what the clinician does with the feedback. Two values were computed: Y-OQ *completion* and clinician *viewing* rates for parent and youth respondents. Completion and viewing rates for each patient were then multiplied together and then by 100, to transform the value to a percentage (Bickman et al., 2016). Table 1 describes how to interpret the Implementation Index derived by the two values.

## Planned Analyses

**PRELIMINARY ANALYSES.** Correlations (*r*) were used to investigate associations between study measures. To assess differences on demographic data, independent samples *t* tests were used. For patients in the MFS condition (Wave B), symptom status from pre- to post-treatment was

summarized with the reliable change index (RCI; Jacobson & Truax, 1991). The RCI indicates the smallest change between two scores that can be considered distinguishable from measurement error, by multiplying the standard error of the difference between the two scores. RCI for the current sample was computed to be 10 points. *Deterioration* status is reflected when a patient worsens by at least 10 points compared to patient's baseline score. *No reliable change* status reflects a change of less than 10 points. *Improvement* reflects patient improvement by at least 10 points, but patient's final score remains above the clinical cutoff score (total score of 29). *Recovery* status reflects a similar improvement of at least 10 points, but the final score falls below the clinical threshold (Jacobson & Truax, 1991).

**MAIN ANALYSES.** Hierarchical linear modeling (HLM) was used to examine the current study's empirical aims. **TO MEASURE DROPOUT AND TREATMENT LENGTH, TWO SEPARATE MULTILEVEL MODELS WERE USED WITH PATIENTS NESTED WITHIN CLINICIANS.** To measure rate of change in outcome scores across time for patients in Wave B only, growth curve modeling with a three-level repeated measures design was used **WITH MEASUREMENT OCCASIONS NESTED WITHIN PATIENT NESTED WITHIN CLINICIANS. THE CHOICE TO USE HLM TO TEST BOTH OF THE STUDY'S HYPOTHESES WAS INFORMED BY THE TYPE OF DATA COLLECTED AND THE INCLUSION OF RANDOM FACTORS. PATIENTS AND CLINICIANS REPRESENT RANDOM FACTORS, DUE TO THE FACT THAT IF THE STUDY WAS REPLICATED, THE NEW SAMPLE WOULD INCLUDE DIFFERENT SETS OF PATIENTS AND CLINICIANS (PITUCH & STEVENS, 2015). MFS** effects were conceptualized to correlate across clinician and clinicians' caseloads of patients. Also, the final data set included an uneven number of patients and clinicians per condition, meaning sample sizes across waves differed as well as number of patients assigned to each clinician's caseload. Lastly, the number of Y-OQs administered and collected from each patient case differed. These factors in addition to

obtaining a better understanding of factors that affect outcome on the basis of a number of independent variables, including a time variable (i.e., measurement occasions), contributed to the choice to use HLM (Raudenbush, & Bryk, 2002). The statistical package, HLM7 (Raudenbush, Bryk, Cheong, Congdon, & Du Toit, 2011), was used to run the main analyses.

***Hypothesis 1a.*** The use of an MFS will decrease dropout rate, and this relationship will be moderated by a patient's urgency at intake.

***Analysis for hypothesis 1a.*** Dropout is a dichotomous dependent variable; therefore, a hierarchical generalized linear model, specifically a logistic multilevel regression, was chosen to examine whether implementation of an MFS affected probability of dropout (the odds of $Y = 1$). Logistic regression models include a specific binomial *Bernoulli* distribution for the dependent variable, a linear regression equation, and a link function [$\text{logit}(p) = \ln(p/(1-p))$]. The last component transforms the dichotomous outcome variable into a continuous one called the natural log of the odds of $Y = 1$, which can be interpreted as an odds ratio for dropout for those who sought treatment before or after the MFS was implemented (Pituch & Stevens, 2015). The parameters in these equations were not provided in terms of the outcome variable, but of the underlying variate that is on a scale ranging from $-\infty$ to $+\infty$, defined by the logit transformation of the $\text{logit}(p) = \ln(p/(1-p))$. The logistic function transforms these predictions back into values between 0 and 1, which can then be interpreted as predicted probabilities of dropout according to each predictor.

Within the logistic regression equation, dropout was coded as 1 and treatment completion (i.e., non-dropout) was coded as 0. Patient-level explanatory variables included condition ("wave"; Wave A = 0, Wave B = 1) and urgency at intake ("urgency"; i.e. proxy for severity at baseline). Variability between clinicians in patient-level urgency of intake was computed by averaging urgency by intake scores of each clinician's caseload, representing caseload severity. Both

predictors were group mean centered and clinician variability across urgency of intake ("meanurgency$_j$") was grand mean centered. Additional patient-level variables, ethnicity and diagnostic profile, were examined to assess whether they explained additional variance in dropout, as they have in past studies (De Haan et al., 2013). Both ethnicity and diagnostic profile are dichotomous variables and were entered uncentered. Model parameters were estimated using the full maximum likelihood procedure. Explanatory variables were added level by level and retained based on each predictor's significance test. Variance accounted for by the grouping variable, or clinician, was measured with the intraclass correlation. The unconditional logistic regression equation is provided below.

Level 1:

$$\ln(\text{odds } Y = 1) = \beta_{0j} + r_{ij}$$

Level 2:

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

For illustration purposes, the proposed conditional model based on research question 1a with all predictors entered is as follows.

Level 1:

$$\ln(\text{odds } Y = 1) = \beta_{0j} + \beta_{1j}(\text{urgency}_{ij} - \overline{\text{urgency}}_j) + \beta_{2j}(\text{wave}_{ij} - \overline{\text{wave}}_j) + r_{ij}$$

Level 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{meanurgency}_j) + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}(\text{meanurgency}_j) + u_{1j}$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}(\text{meanurgency}_j) + u_{2j}$$

The outcome, ln (odds $Y = 1$), represents a patient's dropout status for a specific clinician. The regression coefficient, $\beta_{0j}$, is the mean dropout score at level 1 and transforms to an outcome

variable predicted by meanurgency$_j$ when elevated to level 2. $\beta_{1j}$ is the within-patient association between patient urgency and dropout, then as predicted by meanurgency$_j$ at level 2. $\beta_{2j}$ represents the impact of wave (i.e., "MFS's effect") on dropout, the relationship depending upon clinician meanurgency$_j$ at level 2. The fixed effect, $\gamma_{00}$, is the average of patient-level intercepts across clinicians, or the dropout mean. $\gamma_{01}$ is main effect of meanurgency$_j$ on dropout, holding wave constant. Another fixed effect, $\gamma_{10}$, is the average of the patient-level urgency and dropout slopes across clinicians, holding wave constant. $\gamma_{11}$ is an interaction variable, and if it is positive and significant, the effect of patient-level urgency on dropout is larger for those clinicians who serve patients with higher levels of symptom severity. $\gamma_{20}$ represents the difference between the two waves on dropout rate, holding urgency constant. $\gamma_{21}$ is another interaction variable that examines whether the wave effect depends on meanurgency$_j$. At the group level, $u_{0j}$, $u_{1j}$ and $u_{2j}$ are all level-2 residuals, assumed to have means of zero and be independent from the residual errors at the patient level.

***Hypothesis 1b.*** The use of an MFS will increase treatment length per treatment episode, and this relationship will be moderated by urgency at intake.

***Analysis for hypothesis 1b.*** For 1b, another two-level multilevel model was used to examine where implementation of an MFS impacted treatment length. Patient-level variables (level 1) were nested within clinicians (level 2). Explanatory variables included condition ("wave"; Wave A = 0, Wave B = 1) and urgency at intake ("urgency"; i.e., proxy for severity at baseline). Urgency at intake for each of one clinician's patients was averaged ("meanurgency$_j$"), serving as a clinician-level variable to account for clinician caseload severity in treatment length, and whether urgency-sessions slopes vary across clinicians. Patient-level predictors were group mean centered,

and the clinician-level variable was grand mean centered. Full maximum likelihood was used to estimate unconditional and conditional models through chi-square tests of significance. Variance accounted for by the grouping variable, or clinician, was measured with the intraclass correlation. The unconditional model is provided below.

Level 1:

$$Y_{ij} = \beta_{0j} + r_{ij}$$

Level 2:

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

For illustration purposes, the linear model fitted with the theorized individual- and clinician-level predictors is provided here:

Level 1:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(\text{urgency}_{ij} - \overline{\text{urgency}_j}) + \beta_{2j}(\text{wave}_{ij} - \overline{\text{wave}_j}) + r_{ij}$$

Level 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{meanurgency}_j) + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}(\text{meanurgency}_j) + u_{1j}$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}(\text{meanurgency}_j) + u_{2j}$$

$Y_{ij}$, is the outcome, a patient *i*'s treatment length for clinician *j*. The regression coefficient, $\beta_{0j}$, is the mean treatment length at level 1, becoming an outcome variable in level 2 predicted by $\text{meanurgency}_j$. $\beta_{1j}$ is the within-patient association between urgency at intake and treatment length, controlling for wave, then as predicted by $\text{meanurgency}_j$ at level 2. $\beta_{2j}$ represents the wave impact (i.e., "MFS's effect") on treatment length, controlling for urgency at intake, and then the relationship becomes dependent upon clinician $\text{meanurgency}_j$ at level 2. The fixed effect, $\gamma_{00}$, is

the average of patient-level intercepts across clinicians, or the session mean. $\gamma_{01}$ is main effect of $meanurgency_j$ on treatment length. Another fixed effect, $\gamma_{10}$, is the average of the patient-level urgency and treatment length slopes across clinicians. $\gamma_{11}$ is an interaction variable, and if it is positive and significant, the effect of patient-level urgency on treatment length is larger for clinicians who serve patients with higher levels of symptom severity. $\gamma_{20}$ represents the difference between two waves on treatment length. $\gamma_{21}$ is another interaction variable that examines whether effect of wave depends on $meanurgency_j$. The random effect, $r_{ij,,}$ is the within-patient residual and is normally distributed with a mean of zero and variance equal to $\sigma_2$. At the group level, the random residual error term, $u_{0j}$, is the deviation of a given clinician's treatment length mean from the overall average of treatment lengths. $u_{1j}$ and $u_{2j}$ depict clinician variability around their respective parameters, also assumed to have means of zero and be independent from patient-level residual error.

***Hypothesis 2.*** Patients whose clinicians access their feedback more often will improve at a faster rate than patients whose clinicians access feedback less often, and this relationship will be moderated by urgency at intake.

***Analysis for hypothesis 2.*** A three-level growth curve model assessed whether intervention dosage ("intdos"; measured by the Implementation Index) and patient severity at intake ("urgency") affected rate of symptom change for patients in Wave B. Outcomes include repeated measurement of symptoms reported by parent and youth; therefore, two identical models were designated for each reporter. Time was nested within patients (level 1), and patients (level 2) were nested within clinicians (level 3). Weeks from intake versus session number was used as the measure for Time based on previous models of best fit (Cannon et al., 2010). Time and patient-level variables were group-mean centered, and the clinician-level variable was grand-mean

centered. Time was group-mean centered because patients did not begin treatment at the same time but at various times throughout the year, causing each time point to be specific to each case (e.g., time point 2 did not represent another patient's time point 2), causing the intercept to represent average mid-treatment scores. Since only patients from Wave B were included, mean improvement across treatment episodes cannot be attributed to the MFS. The slope coefficient representing the interaction between time and urgency at intake and intervention dosage is of main interest in the following analyses. For any significant interaction, simple slopes were examined for directionality.

For both models, unconditional growth curve models that included time were first fitted to investigate whether symptoms changed significantly over treatment episodes. Inspection of individual growth curves informed whether time should be squared to reflect a polynomial function. Then, urgency at intake, intervention dosage, and a third interaction variable of "urgencyXintdose" were added at level 2. The interaction variable was constructed after group-mean centering urgency at intake and intervention dosage and then multiplying these two numbers together (Aiken & West, 1991). The interaction variable was included uncentered, and urgency at intake and intervention dosage were group mean centered. Since membership overlapped across both waves for some clinicians, a control variable, clinician "cohort," was included at level 3 (dummy coded 0 if hired before MFS implementation or during Wave A, 1 if hired afterward implementation or during Wave B). The unconditional model that follows represents the model fitted to both parent- and youth-reported outcome data.

Level 1:

$$Y_{tij} = \pi_{0ij} + \pi_{1ij}time_{tij} + e_{tij}$$

Level 2:

$$\pi_{0ij} = \beta_{00j} + r_{0ij}$$

Level 3:

$$\beta_{00j} = \gamma_{000} + u_{00j}$$

The proposed conditional growth curve model with individual- and clinician-level predictors entered is provided below. The same structure was proposed for parent- and patient self-reported symptoms.

Level 1:

$$Y_{tij} = \pi_{0ij} + \pi_{1ij}time_{tij} + e_{tij}$$

Level 2:

$$\pi_{0ij} = \beta_{00j} + \beta_{01}(urgency_{ij} - \overline{urgency}_j)_{1ij} + \beta_{02}urgencyXintdose_{2ij} + r_{0ij}$$

$$\pi_{1ij} = \beta_{10j} + \beta_{11}(urgency_{ij} - \overline{urgency}_j)_{1ij} + \beta_{12}(intdose_{ij} - \overline{intdose}_j)_{2ij} + \beta_{13}urgencyXintdose_{2ij} + r_{1ij}$$

Level 3:

$$\beta_{00j} = \gamma_{000} + \gamma_{001}(cohort)_j + u_{00j}$$

$$\beta_{01j} = \gamma_{010} + \gamma_{011}(cohort)_j + u_{01j}$$

$$\beta_{02j} = \gamma_{020} + \gamma_{012}(cohort)_j + u_{02j}$$

$$\beta_{10j} = \gamma_{100} + \gamma_{101}(cohort)_j + u_{10j}$$

$$\beta_{11j} = \gamma_{110} + \gamma_{111}(cohort)_j + u_{11j}$$

$$\beta_{12j} = \gamma_{120} + \gamma_{121}(cohort)_j + u_{12j}$$

$$\beta_{13j} = \gamma_{130} + \gamma_{131}(cohort)_j + u_{13j}$$

Assuming linear growth over one year's time, $Y_{tij}$ is the outcome of patient *i*, of clinician *j*. Also at level 1, the intercept, $\pi_{0ij}$, is a patient's outcome score when time is equal to the average number of weeks a patient has been in treatment. $\pi_{1ij}$ is the model slope and the linear growth rate of the outcome. At level 2, $\beta_{00j}$ represents the mean outcome score across patients, becoming

conditional on the other covariates at the next level. $\beta_{01j}$ and $\beta_{02j}$ are the unique effect of urgency at intake and the interaction variable on mean outcome status, respectively. $\beta_{10j}$ is mean symptom change across patients taking into account all of the covariates. $\beta_{11j}$ is the effect of urgency at intake on symptom change for patient with an average intervention dose, and $\beta_{12j}$ is the unique effect of intervention dose on symptom change for patients with average urgency at intake. $\beta_{13j}$ answers whether the effect of intervention on symptom change depends upon change across urgency levels. The model residual at level 1 is denoted as $e_{ti}$ and has a mean of zero and variance equal to $\sigma^2$. The residuals at level 2 represent the variation of the individual $\pi$s around the means, and the variances of these distributions are denoted as $\sigma_{002}$, for example. Level 3 coefficients and residual terms will be defined if enough variance exists that could be explained by inclusion of the third level.

Chapter 4

# RESULTS

## Preliminary Analyses

A sample of 538 youth who were enrolled into individual therapy services between February 2016 and February 2018 were included in the following analyses. Data were evaluated and found to be within normal limits with regards to outliers and degree of normality. Urgency at intake and youth-reported mean Y-OQ scores were not found to significantly correlate with each other, $r$ = -.001, $p$ = .974. Urgency at intake and parent-reported mean Y-OQ scores did significantly correlate, $r$ = .159, $p$ < .001. Independent t-tests revealed some differences in patient characteristics between Wave A and Wave B, which are detailed in Tables 2 and 3. Patients in Wave B exhibited significantly higher urgency at intake compared to patients who enrolled into services the year prior. Higher mean urgency at intake scores were also reflected in clinicians' caseloads (Table 3).

For patients in Wave B, symptom severity ratings (Y-OQ data) were available for 204 patients (79% of all patients included in Wave B). Of these 204 patients, questionnaire data were collected 82% and 84% of the time before an individual session from a parent and/or a patient, respectively, referred to as the "completion rate." Y-OQ scores collected at baseline were compared to ratings collected at the last session in order to assess reliable change across treatment episode. According to parent-report of patient's symptoms ($n$ = 203), 52% of the sample did not exhibit reliable change, 15% deteriorated, 15% improved, and 19% recovered, or exhibited positive reliable change, from pre- to post-treatment (or before dropout). According to youth-report ($n$ = 103), 54% did not experience reliable change, 14% deteriorated, 11% improved, and 20% recovered. Findings are similar to what has been found in the past for youth served in

community mental health settings (Smith & Jensen-Doss, 2017; Warren et al., 2012), though the sample was less severe at intake compared to other samples in the same type of setting (Brown & Jones, 2005; Cannon et al., 2010). Of note, 42% of patients began treatment in the "non-clinical" range of the measure according to parent-report, 41% according to youth-report.

## Main Analyses

**DROPOUT RATE.** On average, dropout rate for all patients across both waves was approximately 45%. Forty-four percent of patients dropped out of treatment in Wave A; 46% of patients dropped out in Wave B. Intraclass correlation (ICC) calculations indicated that an estimated 7.48% of the overall variance in the model could be accounted for by differences between clinicians. Due to a non-significant test results for $u_{1j}$, predictors were fitted to a random intercept model. Variance in predictor-session slopes across clinicians had to be held constant; therefore, covariance of means and slopes could not be obtained. Dropout was not found to differ across waves ($B$ = 0.01, $p$ > .05, CI: .701, 1.454), nor did level of urgency at intake help to explain a significant portion of variance in dropout ($B$ = 0.11, $p$ > .05, CI: 0.940, 1.321). Due to these results, the level 2 predictor (meanurgency) was not added to the model.

Following De Haan and colleagues' (2013) recommendation, diagnostic profile and race were added together into the unconditional model to assess whether they could explain any variance in dropout rate for the sample as a whole. Diagnosis type was included as a dichotomous variable, distinguishing between patients with an internalizing diagnostic profile (e.g., disorders related to anxiety and depression, coded as 0) and patients with an externalizing diagnostic profile (e.g., disorders related to behavioral dysregulation, coded as 1). Race was included as a count variable that included frequencies of racial identity collected via parent report at intake. White/Caucasian patients ($n$ = 218) were coded as 1, Latinx/Hispanic patients ($n$ = 277) as 2,

African American patients ($n$ = 42) as 3, Asian American patients ($n$ = 3) as 4, and patients who indicated "Other" ($n$ = 5) as 5.

Diagnostic profile type significantly explained variance in dropout rate for the entire sample ($B$ = 0.41, $p$ = .033, CI: 1.034, 2.207), controlling for race. For patients with an externalizing diagnostic profile, the odds ratio for dropout was 1.54 times the odds of internalizing patients dropping out of treatment. Probabilities of dropout for externalizing and internalizing patients were estimated to be 50% and 42%. Additionally, race/ethnicity was found to significantly account for unique variance in dropout rate ($B$ = 0.43, $p$ < .001, CI: 1.227, 1.942), above the variance accounted for by diagnostic profile. Differences in dropout rate amongst White/non-Hispanic ($n$ = 218) and Latinx/Hispanic ($n$ = 277) patients were examined in isolation due to the low number of patients represented within the other racial/ethnic groups. Latinx/Hispanic patients exhibited an odds of dropout 1.38 times the odds of dropout for White/non-Hispanic patients. Probabilities of dropout for Latinx/Hispanic and White/non-Hispanic families, were estimated to be 46% and 38%, respectively. See Figure 6 for an illustration.

**TREATMENT LENGTH ACROSS WAVES.** On average, treatment length for all patients in Wave A and B was 8.54 ($SD$ = 6.29) sessions that spanned across approximately 14 weeks ($M$ = 14.06, $SD$ = 8.58). ICC was calculated from two unconditional models to represent the two measurements of treatment length (number of sessions and time in days from intake). Calculations estimated 4.72% and 2.93% of the overall variance in number of sessions and days in treatment, respectively, could be accounted for by differences between clinicians. Since variance was greater, treatment length measured through number of sessions was used as the outcome variable. Due to a non-significant test results for $u_{1j}$, predictors could only be fitted to a random intercept model. Variance in predictor-session slopes across clinicians had to be held constant; therefore, covariance

of means and slopes could not be obtained. Treatment length was not found to differ across waves ($B$ = -0.05, $SE$ = 0.05, $p$ > .05), nor was level of urgency at intake found to explain any of the variance in treatment length ($B$ = 0.13, $SE$ = 0.15, $p$ > .05). Due to these results, the level 2 predictor (meanurgency) was not added to the model.

**OUTCOME DATA WITHIN WAVE B.** Two 3-level growth curve models were fitted to parent- and youth-reported symptom severity, collected over time within the Wave B cohort. Clinician effects were found to account for less than .001 of the variability within both models, indicating that the data were better represented by two-level models where time was nested within patients (Schiefele et al. 2017). ICC calculations indicated that an estimated 75.47% and 83.54% of the overall variance in parent- and youth-reported outcomes, respectively, could be accounted for by differences between patients. See Table 4 for results of the hypothesized models. In an attempt to explain more of the individual-level variance in both models, additional patient time-invariant factors were examined along with significant hypothesized predictors. Table 5 includes both hypothesized and exploratory predictors within one model. Table 6 illustrates the final, best-fit conditional models for both parent- and youth-reported outcome data collected across treatment episode.

Parent-reported Y-OQ scores. Examination of individual growth curves indicated a polynomial function, and Time$^2$ was added to the model to capture the equation's quadratic form. Time and Time$^2$ explained 35.14% of the variance in scores. On average, symptom scores over time decreased each week by approximately 0.51 of a point ($SE$ = 0.19, $p$ = .007). Patients with an average urgency at intake score ($M$ = 4.17, $SD$ = 0.77) also had an average Y-OQ score of 30.02 across treatment episode, which is above the suggested clinical cutoff of 29 (Burlingame et al., 2004). For every unit increase in a patient's urgency at intake score, patient's average Y-OQ score

increased by 2 points (*SE* = 0.79, *p* = .012). Urgency at intake did not predict change of Y-OQ over time (*B* = -0.091, *SE* = 0.065, *p* > .05), as hypothesized. The polynomial function of time ($Time^2$) estimated that rate of change slowed toward end of treatment (*B* = 0.01, *SE* = 0.005, *p* = .024).

The MFS intervention dosage was found to impact rate of change significantly over time, as every point increase in the Implementation Index, rate of change further increased by 0.006 of a point (*SE* = 0.003, *p* = .035), on the Y-OQ. The Implementation Index did not significantly interact with urgency at intake (*B* = 0.02, *SE* = 0.03, *p* > .05), as hypothesized. Symptom severity decreased faster for those who received a higher dose of the intervention compared to symptom change for the sample as a whole. For Y-OQ completion rate, parents were administered and completed a questionnaire immediately before session on average 81.89% of the time. Then from the questionnaire data available, clinicians viewed parent feedback before the next session 58.72% of the time, on average. Taken together, the Implementation Index, or the intervention dose, equaled a mean of 47.61% across all clinicians and all parents who provided feedback at least once during a treatment episode. See Figure 7 for change over time for intervention dosage separated by urgency at intake scores.

Less than 1% of the variance in Y-OQ scores was explained by Implementation Index and urgency at intake (PVE = .0025). Time-invariant patient-level variables were then added to the model to evaluate whether they could explain additional variance in parent-reported Y-OQ scores. For patients whose score trajectory classified them as “not-on-track” (NOT), an alert was released to their clinician in real time. For patients who met NOT criteria, their Y-OQ scores were estimated to be on average 10 points higher than “on-track” patients’ mean Y-OQ score (*B* = 9.52, *SE* = 2.20, *p* < .001). Diagnosis type was also included as a dichotomous variable, distinguishing between

patients with an internalizing diagnostic profile (e.g., anxiety, coded as 0) and patients with an externalizing diagnostic profile (e.g., behavioral problems, coded as 1). Patients with primarily externalizing problems exhibited a flatter slope, or a slower rate of change ($B$ = .37, $SE$ = 0.18, $p$ = .039), compared to the patients diagnosed with primarily internalizing problems. See Figure 8 for illustration of growth rates separated by diagnosis type.

All patient-level predictors included in the final model were estimated to explain 11% of the variance in parent-reported outcomes (not including Time or Time$^2$). Effect size (ES) related to change over time was calculated by taking the difference between the standard deviations of the growth parameters. Through this measurement, Implementation Index and diagnostic profile exhibited a very small effect on symptom reduction over time (ES = .022). Cohen's $d$ was also calculated by examining mean differences in Y-OQ scores after being in treatment for 4-6 weeks between high and low Implementation Index scores (i.e., sample split at 50% implementation). Difference in means was divided by the pooled standard deviation of the two groups, resulting in an estimate of .029 (Cohen, 1988).

Self-reported Y-OQ scores. Growth variance was found to be significant and linear, explaining 43.60% of the variance in outcome scores. On average, patients began treatment at a Y-OQ score of 37.12, above the clinical cutoff of 29 (Burlingame et al., 2004). Scores decreased 0.61 of a point each week ($SE$ = 0.12, $p$ < .001) across patients. For computing intervention dosage, youth were administered and completed a questionnaire immediately before session on average 83.73% of the time. From the questionnaires available, clinicians viewed youth feedback before the next session 60.51% of the time, on average. Taken together, the Implementation Index for youth report of symptom severity, or the intervention dose, equaled a mean of 51.09% across all clinicians. Neither MFS intervention dosage ($B$ = -0.01, $SE$ = 0.01, $p$ > .05), nor urgency at intake

($B$ = 0.61, $SE$ = 1.51, $p$ > .05), was found to significantly predict symptom change or symptom severity at one point in time, respectively. Urgency at intake also did not significantly predict change over time in Y-OQ score, ($B$ = -0.08, $SE$ = 0.10, $p$ > .05), as hypothesized. Diagnostic profile was included and found to significantly explain some of the variance in youth-reported symptom severity (PVE = .0226). On average, patients with an externalizing diagnostic profile were found to have a mean Y-OQ score that was 9 points ($SE$ = 3.13, $p$ = .005) below the average score of patients with an internalizing diagnostic profile. Youth with externalizing problems reported fewer symptoms than internalizing youth. See Figure 9 for an illustration.

Chapter 5

# DISCUSSION

The current study joins a handful of other studies that have examined the use of a measurement and feedback system (MFS) within the context of mental health services for youth. An MFS is online platform that represents the physical vehicle for carrying out the evidence-based strategy of progress monitoring and feedback. Two steps contribute to an MFS's function: data collected from patients and clinician's access to their patients' data. The current study included a historical, non-equivalent control design that also served as a program evaluation of an MFS, newly implemented by one community mental health clinic (CMHC) serving youth and families. The first aim of this investigation was to assess whether the agency-wide rollout had an effect on patient treatment engagement, i.e., premature termination and treatment length. The second aim was to assess the impact of MFS implementation on the outcomes of patients whose progress reports were collected and then made accessible to their clinicians. Though collection of questionnaires was mandated by the agency, clinicians had the choice to access their patient's feedback and whether they did or not was evaluated for its effect on symptom trajectories while in treatment.

The following sections move from analyses conducted with the entire sample to analyses with a subsample. The entire sample included patients who were enrolled into individual therapy services across two years. At the half-way mark between the two years, an MFS was implemented agency-wide, splitting each year into "Wave A" and "Wave B." Treatment engagement, by all measures, was not found to differ across the two waves, as hypothesized. No longer separated by wave, two patient-level variables, race and diagnostic profile, were found to explain some of the variance in dropout rate. Then, patients in Wave B were included in a second set of analyses. For

a subsample of patients in Wave B, outcome data related to symptom severity according to parent and youth self-report were available. According to parent report, MFS implementation level was found to impact the rate in which patients' symptoms improved.

## Comparison of Waves on Patient Treatment Engagement

**DROPOUT.** A case was determined to have dropped out if two or more missed appointments in a row were documented by the clinician, whether these took place after the first appointment or fourteenth. This decision was made to best reflect premature termination according to the treating clinician and per recommendation of the dropout literature (De Haan et al., 2013). In the current study, almost half of patients across both waves were classified as having dropped out. This estimate is common for the community setting (De Haan et al., 2013). Use of an MFS was hypothesized to impact the dropout rate, based on the assumption that clinicians would be better able to predict and then prevent premature termination once privy to patients' perspective on their own progress (Kendrick et al., 2016; Lambert et al., 2018; Shimokawa et al., 2010). The rate of dropout was not found to be affected by the MFS, including for patients judged to be more severe than others at baseline, as hypothesized. Similarly, a recent meta-analysis that included only youth found insufficient evidence to conclude use of an MFS reduces dropout (Bergman et al., 2018). It is also possible that more time than one year immediately after rollout may be necessary to begin to detect an MFS's effects on dropout (Brattland et al., 2018).

Patient predictors were then examined across the entire sample in an effort to better help explain the dropout rate. Latinx/Hispanic patients were more likely to have dropped out compared to their White/non-Hispanic counterparts. Additionally, families who presented to treatment for behavior problems were more likely to have dropped out compared to families with children who presented to treatment for depression or anxiety. Reasons for dropout in the current study are

unknown, but findings related to ethnicity have been documented in the past (De Haan et al., 2013; Garland et al., 2005; Kapke & Gerdes, 2016). Causal factors were limited by study design, but it is possible the higher rate of dropout for Latinx/Hispanic families is related to access to other resources in the community and barriers caused by various facets of systemic discrimination (Kapke & Gerdes, 2016).

**TREATMENT LENGTH.** The way in which many studies operationalize premature termination varies (De Haan et al., 2013); hence, the current study included a separate set of analyses for treatment length to separate it from dropout. Number of sessions is often of interest due to insurance coverage and number of allotted sessions, but it may not accurately reflect the length of time a patient is on a clinician's caseload (Weisz et al., 2012). This latter factor is important as holding a place on a clinician's caseload often prevents clinicians from being assigned new patients from a waitlist. Treatment length, as measured both by sessions and weeks from intake, was hypothesized to depend upon how "urgent" a patient was judged to be at intake, this study's proxy for patient severity measured in studies that classify patients as either "on-track" (OT) or "not-on-track" (NOT) in reference to their expected trajectories after treatment begins (Kendrick et al., 2016; Lambert et al., 2018). Treatment episodes were expected to be longer for youth with more severe problems at intake since more sessions may be necessary to reduce distress (Gondek et al., 2016; Lambert et al., 2003). Such an effect was not found in the current study, as neither condition nor urgency at intake was found to relate to treatment length.

On average across conditions, treatment episodes lasted approximately three and a half months (14 weeks) and included eight to nine sessions per family. This finding differs from other studies conducted with similar samples with regards to population and setting, as treatment typically lasts on average six to nine months in community settings (Warren et al., 2010; Weisz et

al., 2012), an estimate approximately double the one from the current study. This difference is likely due to the current study's inclusion criteria, which artificially truncated treatment length. Patients who overlapped waves were excluded to enhance the likelihood of preserving the MFS effect if any effect was present to be found. This decision over-selected cases with shorter treatment episodes and filtered out cases with longer treatment episodes as well as patients who enrolled later in the year time frame. The study's selection impinged the study sample's representativeness to the population and overall generalizability. Most importantly, the study's criteria led to a reduction in power, limiting variability within the final sample, and consequentially, lowering the likelihood of finding a true effect of the MFS. These limitations should be kept in mind in review of the following results and are continued further in a latter section.

## MFS effects on Treatment Outcomes in Wave B

**RELIABLE CHANGE.** MFS data collected in Wave B was used to calculate a reliable change index for examining whether a patient's symptoms from pre- to post-treatment differed sufficiently to suggest individual change (Jacobson & Truax, 1991). Collected through the Y-OQ, patients and their parents reported little improvement in symptoms from the beginning to end of treatment. Almost a quarter of youth were classified as "recovered" from intake, though majority of youth experienced no reliable change. These findings may appear lackluster and perhaps discouraging but they are consistent with percentages reported in other studies also conducted in community settings (Boswell et al., 2015; Cannon et al., 2010; Smith & Jensen-Doss, 2017; Warren et al., 2012); these results are not unique to the current study's CMHC. Despite similarity, it is important to consider the impact of study design. As aforementioned, selection criteria most likely reduced variability and the potential for improvement, as more severe scores at intake permit greater change

over time (Smith & Jensen-Doss, 2017). Findings related to pattern of change while in treatment will be discussed in the following sections and provide a more nuanced picture of patient progress, one that cannot be captured through only two time points.

With the measurement caveat in mind, room for improvement in the quality of clinical services delivered to these families remains. The study's CMHC, as well as other clinics that report similar values (Smith & Jensen-Doss, 2017), were part of state- and legislative-mandated training initiatives for incorporating evidence-based treatments (EBTs) into its practices. Meaning, clinicians in the current study had been trained to deliver effective strategies that one could reasonably assume were being deployed in treatment with patients, to some extent. Despite systematic efforts to make treatment more effective, outcomes derived in community settings have not been found to match the level of symptom improvement reported in efficacy studies conducted with the same manualized treatments (Weisz, Uegueto, Cheron, & Herren, 2013; Weisz et al., 2017). Findings from the current study lend additional support that the way in which evidence is being disseminated to real-world clinicians is not working and more must be done to expect greater rates of improvement.

**PARENT-REPORTED OUTCOMES.** Symptom change according to Y-OQ data was found to follow a similar curvilinear pattern that has been described in the past (Cannon et al., 2010; Cooper et al., 2013; Warren et al., 2010), where symptoms tended to improve the most at the beginning of treatment and then level off as treatment continued. Alert status, i.e., how a patient is progressing through treatment, exhibited an effect on Y-OQ score. Those flagged for not making expected progress ("NOT") had significantly higher average Y-OQ scores compared to patients who were making progress as expected ("OT"). Also, urgency at intake predicted a higher average Y-OQ score, as hypothesized. Taken together, a patient who was labeled NOT and exhibited higher

urgency at intake also experienced greater severity in symptoms at one point in time, according to parent report. These relationships are intuitive but lend some confidence in the urgency at intake measure, an item used in the CMHC to assign patients to clinicians. As for change over time while in treatment, diagnostic profile explained a significant proportion of the variance. Parents of children with internalizing concerns reported faster improvement in treatment compared to parents of children with predominantly externalizing concerns. Meaning, children with disruptive behavior exhibited a slower recovery rate. Neither alert status nor urgency at intake predicted trajectory of symptoms over duration of treatment. MFS effects are discussed in the following section.

**YOUTH-REPORTED OUTCOMES.** Multiple perspectives provide unique and important information for practice, especially when a child is experiencing internalizing symptoms (De Los Reyes & Kazdin, 2005). Pattern of youth report has also been found in the past to predict clinical deterioration accurately without including other informants (Cannon et al., 2010). In the current study, youth report of symptom severity was predicted by diagnostic profile. On average, patients with internalizing problems reported greater severity than their parents. Youth diagnosed with a disorder related to behavior problems reported significantly less distress than their internalizing counterparts, an average score that fell below the clinical threshold that warrants treatment, indicating a subclinical level of concern (Burlingame et al., 2004). This latter finding is more characteristic of what has been found in the past, as youth often appear to underreport level of distress when compared to parents (Cannon et al., 2010; De Los Reyes & Kazdin, 2005). Urgency at intake did not predict variance in the intercept or trajectories over time, as hypothesized, and use of the MFS was not found to affect youth report of symptoms at the intercept nor over time.

## Implementation of the Measurement Feedback System

Before addressing results related to impact of the MFS, it is important to consider which families were given the opportunity to provide feedback. Of all patients who received individual therapy services in Wave B, 54% were Latinx/Hispanic and 34% were White/non-Hispanic. Of the families in Wave B who actually received Y-OQ questionnaires to complete, 34% were Latinx/Hispanic and 55% were non-Hispanic/White. The same percentage of African American families was represented in the Y-OQ data subsample, but Hispanic families were underrepresented compared to their proportion within the larger Wave B sample. According to anecdotal reports from clinicians, Hispanic families had difficulty understanding the MFS measure, which was translated by developers (Burlingame et al., 2004). A difference between waves (no MFS vs. MFS) on dropout for these families was not detected; therefore, it is likely that clinicians refrained from collecting data from Hispanic families after deciding it was not feasible or that results would not have been valid. The fact that these families were excluded is unfortunate, especially since progress monitoring and feedback has promise to help adapt effective practice strategies for families who have been underrepresented in efficacy trials (Garland, Lebensohn-Chialvo, Hall, & Cameron, 2017; Ng & Weisz, 2016). Within the same vein, the decision to discontinue progress monitoring is understandable if the clinician is unsure how to interpret its results to influence practice.

**IMPLEMENTATION AND RATE OF CHANGE.** The current study took into consideration recommendations (Bergman et al., 2018; Bickman et al., 2016; Gondek et al., 2016; Kendrick et al., 2016) for assessing clinicians' fidelity to the intervention based on the data available. MFS intervention dosage, defined as percentage that clinicians accessed available feedback via the MFS dashboard, was found to impact rate of change in symptom severity over time according to parent

report. These findings add support that clinician rate of accessing patient feedback matters, and increased access is related to the rate in which patients improve (Bickman et al., 2016). Although use of the MFS was mandated by the CMHC, clinicians in this study, on average, did not access all of the patient feedback made available to them via the MFS dashboard. View rate was observed to be lower than the measure completion rate by approximately 20%, which is not an unexpected finding (Douglas et al, 2015; Warren et al., 2012; Whitcomb, Woodland, & Burlingame, 2018).

**IMPLICATIONS FOR THEORY.** The difference in change trajectories for patients who received a higher dose of the MFS lends more support for some theoretical explanations and not others for *how* outcome monitoring may be effective. Feedback Intervention Theory (Kluger & DeNisi, 1996) and its variates (cf. Gondek et al., 2016; Riemer, Rosof-Williams, & Bickman, 2005; Sapyta et al., 2005) propose that feedback to clinicians becomes effective when the clinician experiences incongruence between their own perception of patient's progress and patient's report. Whether this took place was not measured in the current study; only whether a clinician viewed feedback was captured. Nonetheless, clinician view rate did make up half of the proportion of intervention dosage, and higher amount of views did impact change over time.

**CLINICIAN ATTITUDES AND BEHAVIOR.** These findings illustrate that even though a large system can mandate MFS use into practice, it does not mean all clinicians are practicing measurement-based care (Goldberg et al., 2016; Lyon et al., 2016). Evidence continues to mount that clinician subjective judgement without guidance from standardized or individualized measurement is inconsistent and often, inaccurate (Bar-Kalif et al., 2016; Hannan et al., 2005; Tracey et al., 2014; Treichler & Spaulding, 2018; Walfish et al., 2012). Yet in reality, clinicians choose to put less stock in standardized measurement (Jensen-Doss et al., 2016; Norman, Dean, Hansford, & Ford, 2014) and instead opt to rely on personal judgement alone. Therefore, the

question persists (Bickman, 2008; Sapyta et al., 2005) – How can clinicians be sure that what they are doing in session is helpful? And in some cases, how can clinicians be sure that what they are doing is not harmful? In the very least, clinicians are in the position to provide some explanation concerning treatment gains with low severity patients to rationalize keeping other families with more severe concerns on a waitlist.

Interestingly, surveys conducted with real-world clinicians have found that clinicians are often largely in favor of progress monitoring and assessment. For example, 93% of clinical staff ($n$ = 127) rated the practice strategy as "important to very important" in one study, yet only 16% of the same sample reported having collected data from the same patient more than twice (Batty et al., 2013). Similarly, Bickman and colleagues (2000) surveyed 539 masters-level clinicians and 86% expressed desire to receive patient feedback on a monthly basis, despite 77% having never administered a standardized measure in practice. Jensen-Doss and colleagues (2016) found the majority of masters-level clinicians who responded to a questionnaire (n = 504) exhibited positive attitudes toward measurement and feedback and disagreed with an item that stated these tools would harm alliance with patients. Despite declarative favor, only 62% of these clinicians had ever used an MFS, and 45% preferred to never administer a standardized assessment tool (Jensen-Doss et al., 2016). Of note, clinicians who endorsed a cognitive-behavioral therapeutic orientation were more likely to favor progress monitoring but were not more likely to have used it in their practice (Jensen-Doss et al., 2016).

**REASONS FOR POOR ADHERENCE.** The various reasons why clinicians are hesitant to use progress monitoring to obtain feedback in real practice have been grouped into three interrelated reasons. To help find clarity between the mismatch in clinician attitudes and behavior, Jensen-Doss and colleagues (2017) discovered in a follow-up study that standardized measurement is most

likely at the root of discontent toward MFSs, and that perhaps it has been separated from some clinicians' conceptualizations of progress monitoring. Most often, clinicians believe standardized measures fail to capture what is going on with their patients (Bickman et al., 2000; Jensen-Doss et al., 2016; Norman et al., 2014). Common reports reveal beliefs that these tools depersonalize patients and contain items that are unrepresentative of patients' reported symptoms (Norman et al., 2014). Clinicians believe standardized measurement tools do not provide more information that could be gleaned purely from observation and informal interviewing (Jensen-Doss et al., 2016). Further, clinicians have reported that feedback is not specific enough to lend answers to what should be done in treatment to help patients improve (Batty et al., 2013; Kotte et al., 2016). Specificity and personalization are important to clinicians, who are unsure how standardized measures could provide either (Bickman et al. 2000; Jensen-Doss et al., 2016, 2017).

Second, implementation of a new measure, or an entirely novel system, for providing feedback has been described as too burdensome for an organizational system as a whole (Gleacher et al., 2016). The lack of resources for administrative staff to learn how to administer and then maintain an MFS is thought to be infeasible for community providers without external support (Kotte et al., 2016; Lucock et al., 2015). Some clinicians believe failure of such efforts can also be traced back to families, who are described to find it cumbersome and too time-consuming to complete a measure at each visit (Kotte et al., 2016). Most commonly lamented within this category involves the learning curve experienced by staff and clinicians when using new software or an online platform to collect data in conjunction to using a clinic's existing electronic medical record (EMR) system (Gleacher et al., 2016; Ionita, Fitzpatrick, Tomaro, Chen, & Overington, 2016).

Lastly, clinicians in some studies have acknowledged the lack of adequate training in administration and interpretation of standardized measurement tools and consequently do not know the value it could provide to their practice (Gleacher et al., 2016; Ionita et al., 2016; Whitcomb et al., 2018). Other studies include anecdotal reports from clinicians that reflect this sentiment and the need for more training. For example, the validity of standardized measurement in general has often been questioned, including the belief that mood is primarily reflected in a measure's results and the resulting patient feedback does not detect actual dysfunction (Kotte et al., 2016; Ionita et al., 2016). Consequentially, clinicians report difficulty knowing how to introduce and discuss feedback with patients (Ionita et al., 2016; Kotte et al., 2016; Whitcomb et al., 2018). In summary, clinicians support the concept of progress monitoring and patient feedback and even express desire to implement this strategy in their practice, but barriers related to insufficient support for both staff and clinicians impede delivery.

## Clinical Implications

**TREATMENT EFFICIENCY.** In the current study, the increment added to the symptom change rate over time (.51) by the Implementation Index may appear small at face value (.006), but it must be considered with the rest of the model and other predictors' effects. The mixed model equation may be best represented through a case example, such as a child diagnosed with depression who entered treatment at the average urgency at intake and whose parent's report initiated an alert to the clinician that the patient was not improving at the expected rate (labeled "NOT"). Based on the mixed model equation, this patient would have required 8 weeks in treatment before the parent reported a score that fell below the clinical threshold of 29 – *if* the child's clinician is implementing the MFS at 100%. If the child's clinician partially implemented the MFS at 46%, the level in which clinicians in the study were found to do on average, 12 weeks

would be necessary to expect an equal decrease. If the child's clinician never viewed feedback or if feedback was never collected, the child would require 20 weeks in treatment, or 2 ½ more months than if the clinician fully implemented the MFS with this particular patient.

Time spent in treatment for one family is important to consider in a broader context. Within the current landscape of psychological service delivery, the idea of prolonging care without additional benefit seems inefficient, especially in an economic sense. Between 2011 and 2015, ED visits for youth increased by 28% (Kalb et al., 2019), a higher estimate than the 21% increase documented for the five years prior (Torio et al., 2015), trends unmirrored by the adult population (CDC, 2016). For Hispanic teens only, ED visits for psychiatric reasons rose 91% between 2011 and 2015 (Kalb et al., 2019), with these youth more likely to then need inpatient psychiatric care after presenting to the ED. Data indicate that if youth and families continue to be unable to access mental health services, the rate of presenting to tertiary, more expensive care may continue to increase (DHHS, 2017). Greater efficiency in mental health care means shorter wait times between contacting services and enrolling into treatment, as well as less time to relief once the patient begins treatment (Gondek et al., 2016).

**ACCESS TO CARE.** In the current study, more than a third of patients entered the study under the clinical threshold and did not differ in treatment length from those patients with more severe symptoms. This finding is a meaningful one and must be considered when assessing the capacity of community mental health to serve those in need. Evidence does not indicate that individuals who enter treatment below the threshold of clinical impairment will benefit from therapy or additional sessions (Brown & Jones, 2005). The MFS in the current study, the OQ® Analyst, has a "white" code that is issued once a patient reaches subclinical status, with the sole purpose to signal to the clinician that it is time to consider termination (Lambert et al., 2010). (It is unknown

whether clinicians made use or were aware of these codes.) In actuality, less than one-fourth of youth with psychiatric diagnoses in the United States receive appropriate care from a mental health specialist (Costello et al., 2014). In order to increase the number of children who have the opportunity to take advantage of mental health services, perhaps stakeholders could begin to reexamine and rethink the ways in which community mental health responds to those who present for services in the first place.

Community providers are charged with increasing accessibility and capacity while continuing to rely largely on grant and governmental funding. A way to shorten time in treatment without adding to costs would afford community providers with greater capacity to provide care. Fortunately for agencies, the current study adds to the existing literature that efficiency can be attained through ongoing assessment, versus hiring more providers or overextending the current ones (Bickman, 2008; Bickman & Noser, 1999; Sapyta et al., 2005). Through an MFS's guidance, clinicians could boost efficiency through first assessing which skills are necessary to provide a subclinical patient in order to expect additional gains. Whether the best decision is determined to terminate after a few sessions or continue at a reduced frequency, either increases efficiency. All in all, community centers must acknowledge the fact that access to appropriate services is greatly and disproportionately restricted for some children versus others (Garland et al., 2017; Kapke & Gerdes, 2016); therefore, capacity must be increased while also improving and maintaining quality of services.

## Recommendations for Practice

**PRE-SERVICE TRAINING.** The push toward therapeutic practice informed by evidence is based on the logic that strategies with the most empirical support have the best chance of relieving patients of their distress (APA, 2006; Beck et al., 2014). Under different nomenclature, the tenets

of evidence-based practice and its purpose are more than likely reflected in all health professional training programs, not solely the ones that are geared toward research (Beck et al., 2014; Becker-Haimes et al., 2018). With that being said, the evidence is ever-changing and indicated treatments for specific diagnoses change over time; therefore, all masters- and doctoral-level clinicians are encouraged to receive training on how to understand and apply research (Beck et al., 2014; Becker-Haimes et al., 2018; Garland et al., 2017). For example, clinicians stand to benefit from learning of confirmation bias, where it is most likely to materialize in practice, and how it can be corrected through use of standardized measurement tools (Beck et al., 2014).

Service quality in mental health is influenced by several variables outside of a clinician's immediate control (Bickman & Noser, 1999; Macdonald & Mellor-Clark, 2015). Therefore, pre-service training for professionals who wish to deliver effective services is recommended to include training in short-term outcome monitoring. This would allow burgeoning clinicians to begin to protect against bias in their practice, as well as see the effects of new techniques in practice (APA, 2006; Beck et al., 2014; Becker-Haimes et al., 2018; Waller & Turner, 2016). Through guided use, clinicians learn how to adopt and flexibly adapt effective strategies for patient idiosyncrasies (e.g., comorbidity, cultural diversity), unaccounted for in efficacy trials (Bearman & Weisz, 2015; Beck et al., 2014). Treatment gains are proven to the clinician in real time (Pinner & Kivlighan, 2018). Importantly, this approach also helps indicate if a new strategy is hurting, or having unintended effects (Beck et al., 2014). Used in this way, progress monitoring and feedback is the quintessential technique for personalizing care while also improving service quality.

Curricula that include application of research also prepare clinicians to become keen evaluators of new practice strategies after graduation and during post-service trainings (Beck et al., 2014). For instance, what has been documented in past studies is that when an MFS is

introduced into a community setting, one individual often becomes the point person for all clinicians to consult for questions related to the MFS (Borntrager & Lyon, 2015; Ionita et al., 2016). Researchers have warned against this, as spreading ownership across multiple clinicians and staff members increases the likelihood of a new intervention becoming embedded within an organization's culture (Hall et al., 2014).

Stakeholders in charge of implementation must recognize the variability in knowledge and training of clinicians in standardized measurement. The majority of practitioners in community settings have not been trained in the fields that have traditionally emphasized measurement-based care (e.g., clinical psychology) (Becker-Haimes et al., 2018); meaning, those hoping to implement a novel strategy that includes a heavy focus on measurement must take this into account. In order for successful uptake to happen, ongoing support after initial trainings must be provided in order to ensure thorough understanding and practical application (Beidas, Edmunds, Marcus, & Kendall, 2012; Borntrager & Lyon, 2015; Lyon et al., 2016). Extensive on-the-job opportunities for training and explicit instruction (in addition to the initial workshop training) has been described as necessary for adoption by real clinicians and policy leaders alike (Brattland et al., 2018; Garland et al., 2017; Gleacher et al., 2016; Whitcomb et al., 2018). Perhaps most importantly, an essential piece for MFS implementation is greater support for administrative personnel, who are often in charge of on-site administration procedures (Borntrager & Lyon, 2015; Garland et al., 2017; Gleacher et al., 2016; Hall et al., 2014; Kotte et al., 2016).

**CALL FOR SUPERVISION.** According to theory, clinicians have two decisions when they receive feedback from a patient, either to accept or reject, with the latter most often chosen for unexpected feedback (Kluger & DeNisi, 1996). Clinical supervisors are in the critical position to help clinicians assess the relevancy and meaning of feedback data, as no other mechanism exists

at this time within the practice of psychology built to provide an objective, timely, and ongoing appraisal of a clinician's skills. Delivering mental health services is different from other consumer-driven professions in that the practice of psychology is structured in a way where clinicians can rely on post hoc constructions, or explanations, of what took place in treatment (Tracey et al., 2014). Since these explanations are untestable, the clinician gains a sense of learning through experience as their perspective is filed away uncorrected (Tracey et al., 2014). Supervisors have the potential to be the gatekeepers for catching such normal clinician processes, as well as for normalizing a clinician's anxiety around patient feedback and modeling how to discuss it in session with a patient (Becker-Haimes et al., 2018).

Access to supervisors trained in evidence-based practice and knowledge of progress monitoring is especially crucial within psychology (Beck et al., 2014; Becker-Haimes et al., 2018). One of the many reasons why outcomes from efficacy trials do not transfer to practice has to do with poor adherence to the active practice elements that elicit change (Chorpita et al., 2005; Garland et al., 2010; Weisz et al., 2012). Similar to any other intervention, the intention to use an EBT does not materialize into better quality of care. Fidelity checklists provided by a supervisor have the capacity to ensure clinician adherence to an EBT (Bickman, 2008). Use of fidelity checklists, in combination with an MFS, could help illuminate which practice elements are, in fact, drivers of positive change for a particular patient (Bearman & Weisz, 2015). Over time, supervisors and clinicians would hone the capacity to track clinician effectiveness across different patients, and even lead to a greater understanding of which types of patients a clinician delivers the most effective care (e.g., older versus younger children; internalizing vs. externalizing) (Pinner & Kivlighan, 2018).

Whether supervision impacts a clinician's use of an MFS has not yet been determined in an empirical sense, aside from commentaries written on ways to leverage EBP in general (Pinner & Kivlighan, 2018; Tracey et al., 2014; Worthen & Lambert, 2007). Davidson and colleagues (2017) set out to evaluate whether supervision impacted MFS use through two conditions: supervision with clinicians who received feedback data from a month prior, or supervision without feedback. According to clinicians, patients who demonstrated the most improvement were patients whose feedback was discussed in supervision (Davidson et al., 2017). Another study tracked one CMHC's uptake of an MFS and how it was measured to increase clinicians' therapeutic skills over seven years (Goldberg et al., 2016). Goldberg and colleagues (2016) described a supervisory forum designated strictly for discussing clinicians' most difficult patients (as detected by an MFS). Authors believe the forum, a) created an environment where discomfort related to feedback was discussed openly, and b) served to convey to clinicians that the purpose of feedback is to learn new skills for improving clinical skills (Goldberg et al., 2016). Studies with greater control over methods would help to discern how supervision can best support feedback's impact on patient outcomes.

## Limitations

The current study contributes to the small, though rapidly growing, field of health service technologies (Lyons et al., 2016), but it is not without limitations. Since the sample was drawn from archival records, a very limited amount of design features could be controlled, impacting overall power and confidence in the findings. The study used a historical control group for testing the first two hypotheses, but a control group for the second set of hypotheses for treatment outcomes was not available. Special design features must be taken into consideration when

interpreting the study's results, and findings cannot be compared on the same plane to results from an experimental design study that included a real control condition.

Two decisions greatly impacted the sample's representativeness. Inclusion criteria selected patients who began and exited from treatment within one year, in order to prevent treatment episodes from straddling the time in which feedback to clinicians began. This decision was made to increase internal validity of an archival study, but it also most likely inflated dropout rate and excluded more severe patients from the study's sample (De Jong et al., 2014). Excluding patients with longer treatment episodes limited variability within the sample, including more severe patients. Fewer severe patients affected estimates of reliable change, as more severe patients are more likely to attain reliable change and improvement at end of treatment when included (Cannon et al., 2010; Janse et al., 2017; Smith & Jensen-Doss, 2017). Second, the majority of patients met criteria for more internalizing diagnoses than externalizing, when externalizing diagnoses are more commonly represented in community care (Costello et al., 2003; Hawley & Weisz, 2002; Okamura, Jackson, & Nakamura, 2019). A lower proportion of externalizing patients could be due to a subset of families excluded from data collection because of enrollment into two EBTs running at that time, one for young children with disruptive behavior and the other a parenting group for older children. The parenting group targeted behavior management offered either in conjunction with individual therapy for the child or not. Data were not collected from parents who chose to attend group only. Selection criteria and absence of data from all families who presented for services caused the study's final sample to be unrepresentative of typical CMHC settings.

Clinician effects are important to detect, since the intervention's effects were controlled in part by clinician behavior. Attempts were made to control for differences between conditions through clinician-level predictors (cohort, caseload severity), but analyses indicated insufficient

variance for capturing clinician effects. The study was most likely underpowered and included too few patients within clinicians, rather than without differences across clinicians that influenced the type of care delivered (Delgadillo et al., 2018; Goldberg et al., 2016; Schiefele et al. 2017), especially since clinician membership was impacted by the MFS rollout (clinician turnover increased). Additionally, some patients improved faster than others and greater levels of MFS implementation helped to explain some of the variance related to differences in rate of improvement. A likely explanation for this relates perhaps to the type of clinician who was more likely to access patient feedback. These clinicians may have also been more likely to employ other empirically-supported practice strategies (Jensen-Doss et al., 2016), contributing to better outcomes for their patients. Studies that focus on what a clinician does with feedback data have the potential to be greatly influential for practical application and future trainings.

Lastly, the Y-OQ is the only measurement tool that was used in the current study to relay mental health outcome, a multi-faceted construct. The study did not account for family functioning, quality of life, or various other outcomes that capture a youth's functioning with greater comprehension. Additionally, the Implementation Index contains two values, completion rate of the Y-OQ and view rate of the accumulated Y-OQ data (Bickman et al., 2016). It is impossible to know from these values what clinicians may have gleaned from feedback, how it was used, or if it was discussed with supervisors, other clinicians, or the patients themselves. The index therefore provided a proxy of implementation judged by extrapolating one type of clinician behavior. Additional information, perhaps provided by qualitative interviews with clinicians about the nature of their MFS use, might complement this metric. Also, multiple measures for assessing outcome should be included in the future examinations to improve confidence in findings.

## Future Directions

**FOR PRODUCT DEVELOPERS.** Research has moved from targeting single-user adoption of MFS to organizational shifts that help support systemic implementation of progress monitoring. Top-down efforts include finding a way to insert MFS capabilities into a CMHC's existing electronic health record (EHR) system (Lyon et al., 2016). Commercial companies such as the OQ® Analyst have designed their products to fit into pre-existing systems, while others have designed two-in-one EHRs with MFS functionalities. Combining the two systems would ensure administrative staff were trained and constantly exposed to collecting feedback and decrease the burden of having two separate electronic systems that clinicians must access. Systemic changes such as an EHR overhaul may require greater support for administrative personnel (Lyon et al., 2016). Ideally, progress monitoring and feedback would become a default function and part of a clinician's regular workflow, an impossible scenario without greater initial and ongoing support from product developers to administrative personnel.

**FOR IMPLEMENTATION RESEARCHERS.** MFS features that have been found to be more motivating to clinicians over others and lessons learned from past studies could increase buy-in for future MFS applications. Foremost, clinicians are often unaware that the best way to approach feedback data is to combine information gleaned from both normed and idiographic measures (Borntrager & Lyon, 2015; Garland et al., 2017). Meaning, standardized measures with norms plus informal measurement tools constructed with a patient's own words to be used together. Patient-specific information adds specificity to generalized, standardized items, and informs the clinician of a patient's priorities in addition to when a patient may believe treatment can end (Weisz et al., 2011). This recommendation should come as a pleasant surprise to clinicians, as greater specificity in measurement is one of the most common requests when introducing an MFS to practice

(Bickman et al., 2000; Jensen-Doss et al., 2017). When tools include idiographic content, i.e., concerns unique to the patient, clinicians are more likely to be able to form concrete next steps for treatment (Kelley & Bickman, 2009; Sapyta et al., 2005).

When the MFS literature is viewed in its entirety, areas for future research appear limitless as the push for quality assurance continues. What may be the most beneficial for the community of researchers in this area is to make a deliberate effort toward standardizing the language that describes the mechanism(s) being measured within progress monitoring and feedback research. Neglecting to do so has led to various and loosely-related procedures being tested and combined into meta-analytic reviews as if they were the same intervention (Kendrick et al., 2016; Bergman et al., 2018). The consequence of this is reflected by the commercial market, where some "MFS" technologies are being offered to consumers without very basic capabilities, such as feedback on patient status (Lyon et al., 2016). Once researchers adopt a common language, empirical pursuits will become more organized and further aid in detecting change mechanisms (Bergman et al., 2018). These advances would then provide greater clarity to product developers on which features should be optimized in commercial products.

## Conclusion

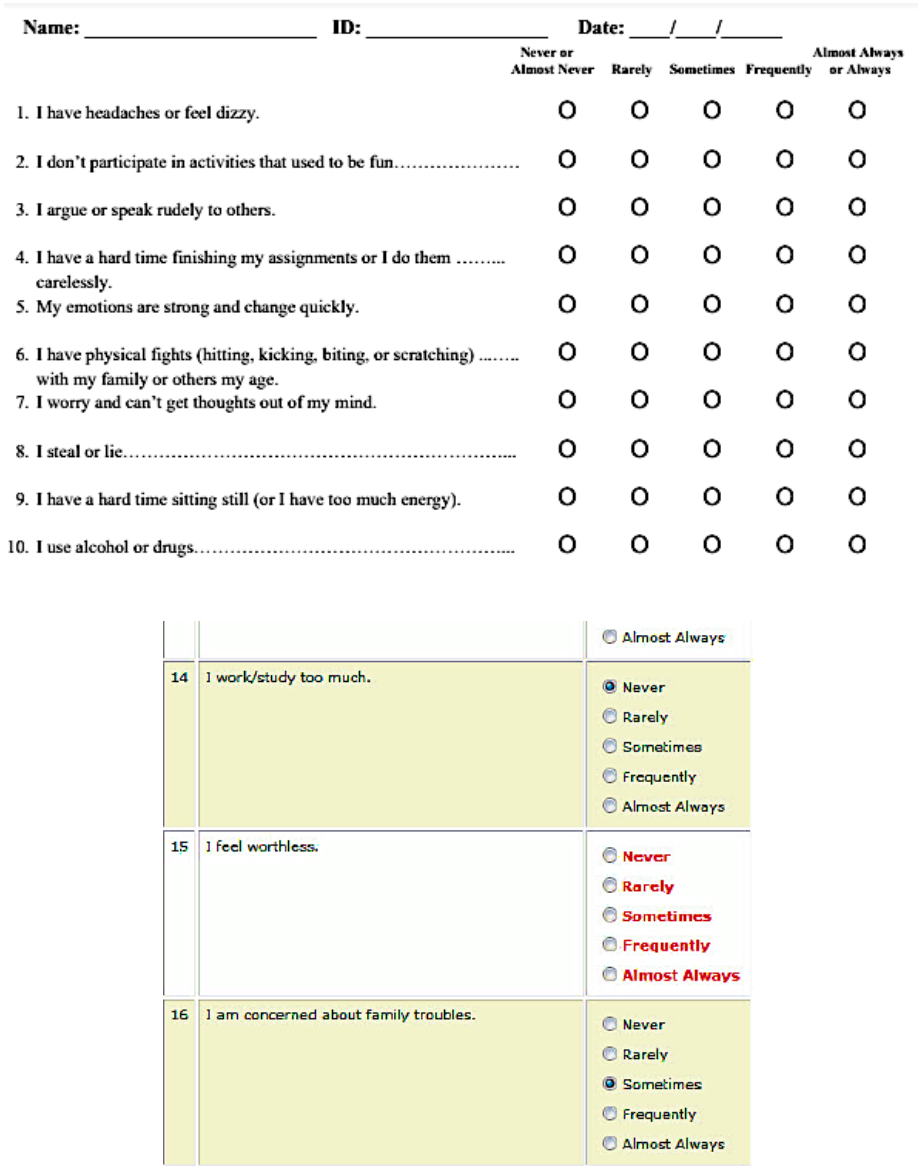
The current study contributes to practice-oriented research by examining the impact of progress monitoring and feedback on the improvement rate of youth who presented for mental health services in one community mental health center. Level of MFS implementation affected the rate in which patients improved while enrolled in treatment, demonstrating the significance that monitoring progress can have regardless of the therapeutic approach or orientation in which a clinician is vested (APA, 2006; Norcross & Wampold, 2011; Worthen & Lambert, 2007). Use of an MFS has the potential to make treatment delivery both more effective and efficient, and greater

treatment efficiency would permit community providers to increase their capacity without increasing costs for care.

Given that effect sizes from meta analyses suggest that the majority of youth with common disorders (e.g., anxiety, disruptive behavior disorders) can show clinical benefit in rigorous RCTs (Weisz et al., 2013, 2017), there remains work to be done to improve outcomes for youth in usual care settings. Many agencies and states have sought trainings in EBTs to improve outcomes, but some studies show that even agencies where EBTs are mandated may not be able to replicate these benefits (Chorpita et al, 2017; Smith & Jensen-Doss, 2017; Weisz et al., 2012). As seen in the current study, the mere introduction of a strategy does not necessarily lead to its effective use. Instead of striving for identical imitation of practice within efficacy studies, clinicians can improve their skillsets by pairing new practice elements supported by the evidence base with progress monitoring tools for ensuring improvement as treatment progresses (Bearman & Weisz, 2015; Ng & Weisz, 2016).

Importantly, authors of the current study acknowledge the difficulty as well as the tumult in providing high-quality therapeutic care to families in crisis. Psychotherapy represents a "noisy" environment, where the common variables (e.g., insurance coverage, transportation) are reliably inconsistent, preventing clinicians from being able to predict what may come next (Macdonald & Mellor-Clark, 2015). Within the same vein, the Ethical Principles of Psychologists and Code of Conduct, Section 2, dictates that psychologists provide services only within boundaries of their competence (APA, 2002). Within the same section, guidelines state psychologists should undertake ongoing efforts to develop and maintain their competence (2.03). Over time, this latter instruction has translated into the formation of institutions that provide continuing education credits for renewal of licensure, equating workshop attendance and online tutorials with

"maintaining competence" without sufficient evidence of long-term gain. Based on the evidence that has accrued, progress monitoring and feedback surpasses the traditional paths for improving practice, while also emerging as a promising apparatus for building competence for delivering effective treatment practices. The time has come to accept what the science has proven both with regards to human error (i.e., clinician bias) and its remedy, ongoing standardized measurement. Then, we can begin to improve the quality of services currently being delivered to families, as well as increase the capacity of the systems that provide them.



**Name:** _____ **ID:** _____ **Date:** ____/____/_____

| | Never or Almost Never | Rarely | Sometimes | Frequently | Almost Always or Always |
|---|---|---|---|---|---|
| 1. I have headaches or feel dizzy. | O | O | O | O | O |
| 2. I don't participate in activities that used to be fun…… | O | O | O | O | O |
| 3. I argue or speak rudely to others. | O | O | O | O | O |
| 4. I have a hard time finishing my assignments or I do them …….. carelessly. | O | O | O | O | O |
| 5. My emotions are strong and change quickly. | O | O | O | O | O |
| 6. I have physical fights (hitting, kicking, biting, or scratching) ……. with my family or others my age. | O | O | O | O | O |
| 7. I worry and can't get thoughts out of my mind. | O | O | O | O | O |
| 8. I steal or lie…….. | O | O | O | O | O |
| 9. I have a hard time sitting still (or I have too much energy). | O | O | O | O | O |
| 10. I use alcohol or drugs…….. | O | O | O | O | O |

| | | |
|---|---|---|
| | | ○ Almost Always |
| 14 | I work/study too much. | ◉ Never<br>○ Rarely<br>○ Sometimes<br>○ Frequently<br>○ Almost Always |
| 15 | I feel worthless. | ○ Never<br>○ Rarely<br>○ Sometimes<br>○ Frequently<br>○ Almost Always |
| 16 | I am concerned about family troubles. | ○ Never<br>○ Rarely<br>◉ Sometimes<br>○ Frequently<br>○ Almost Always |

Figure 1. Client-report measure in hard copy and electronic form. These are two examples of the type of measures and items that are electronically or manually collected from youth and/or parents before each session. Typically, informants are requested to answer the items with the preceding seven to ten days in mind. Responses are consolidated by the MFS and presented to the clinician in similar formats to the one illustrated in Figure 2.

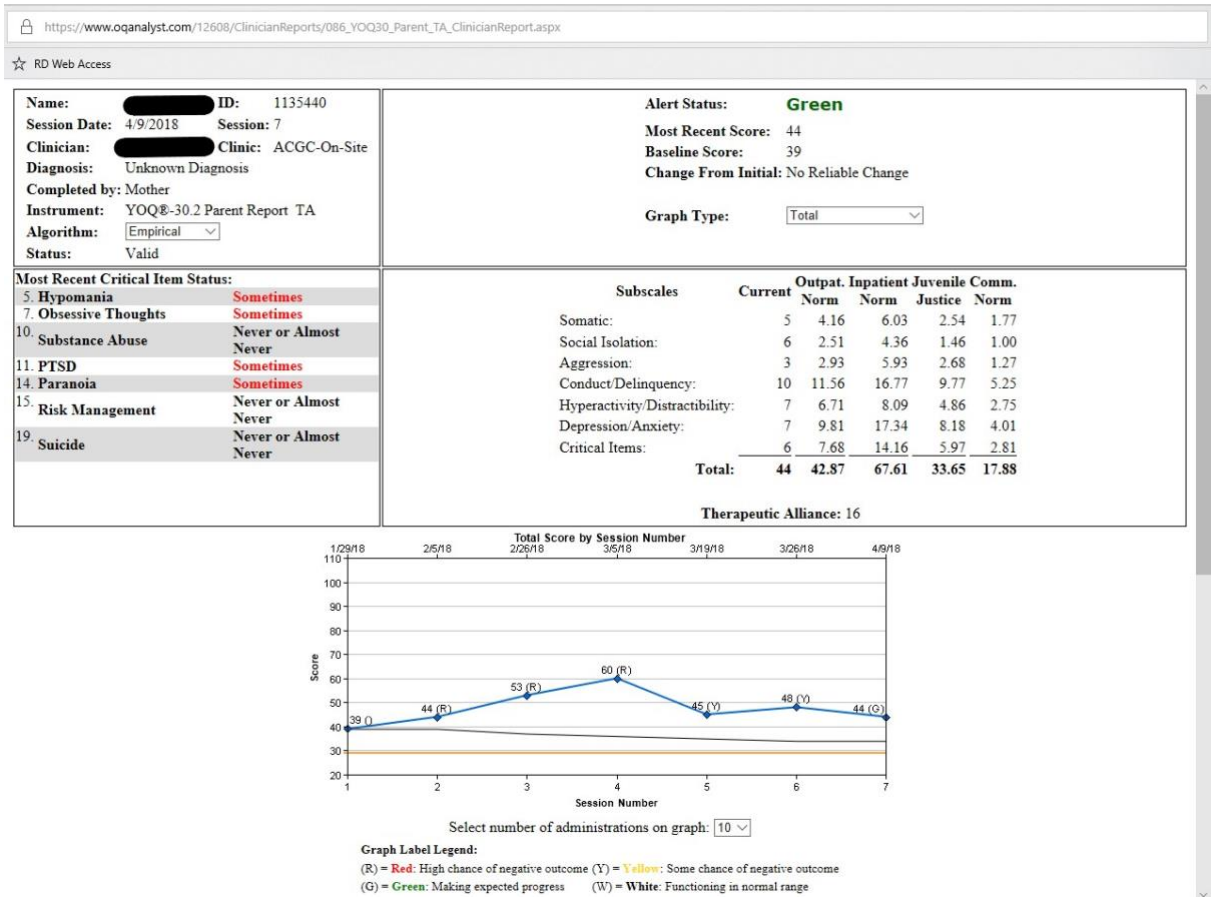



**Figure 2.** OQ® Analyst. This is an example of an MFS dashboard that clinicians access to view patient self-reported data and response to treatment over an episode of care (Lambert et al., 2010). Lines within the graph portion represent clinical cut off thresholds for community and clinical populations. The blue line represents the treatment trajectory of an individual patient and where he/she falls in comparison to same age peers.

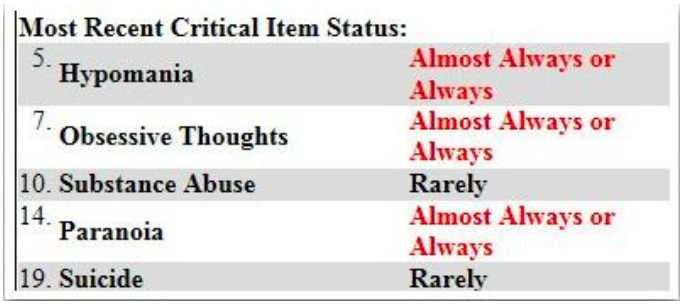



**Figure 3.** Close-up of critical items. This particular MFS uses a patented measure, the Youth Outcome Questionnaire-30 (Burlingame et al., 2004) to measure a youth's progress (Figure 1). Critical items that were most recently endorsed by the parent or youth have been highlighted for the clinician via a section in the MFS dashboard (Figure 2).

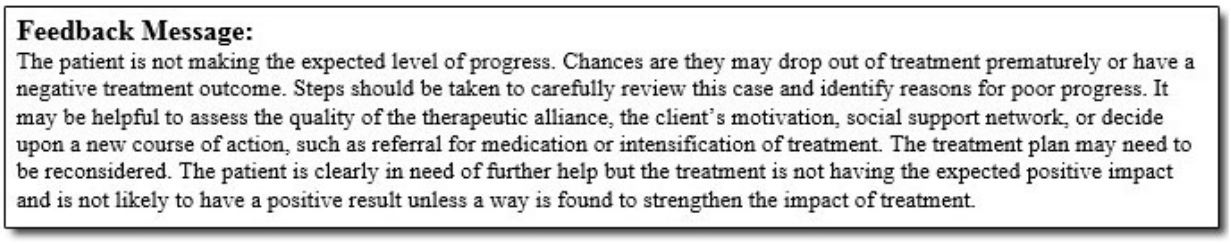



**Figure 4.** Sample feedback message. Underneath measure-specific information from the most recent session and updated progress graph, a feedback message individualized to patient status may be provided.

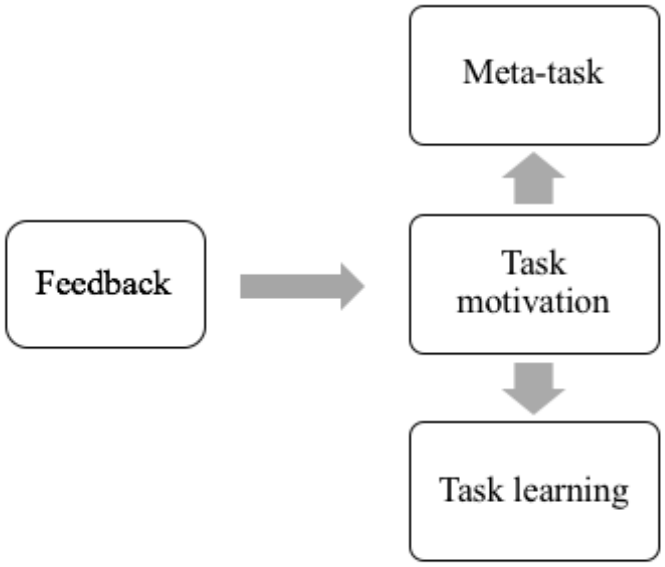



**Figure 5.** Internal levels of control of the Feedback Intervention Theory. This figure is adapted from Kluger & DeNisi (1996).

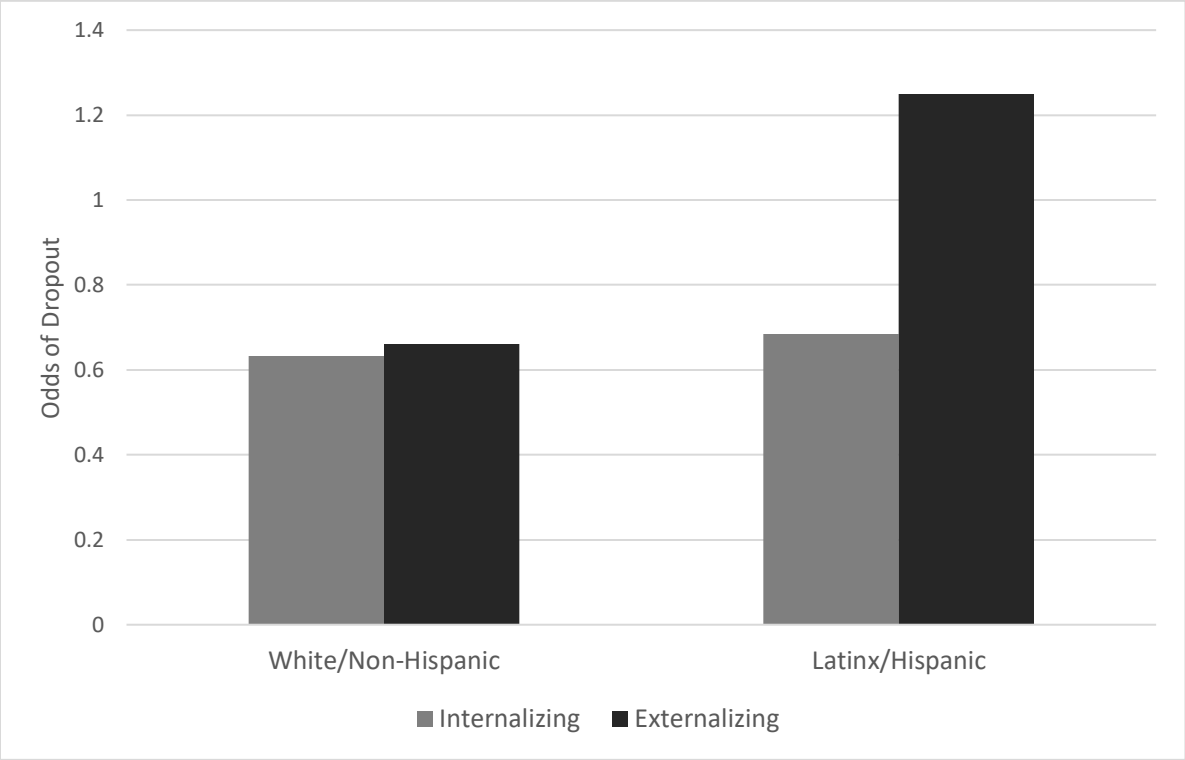



Figure 6. Impact of race/ethnicity on dropout separated by diagnostic profile. The y-axis units are the computed odds of dropout, obtained by dividing the number of patients who dropped out within each subcategory (ex. "White/Non-Hispanic" and "Internalizing") by number of patients who did not drop out. Internalizing problems were less likely to have dropped out compared to those with primarily externalizing problems, and Latinx/Hispanic patients were more likely to have dropped compared to White/non-Hispanic patients.

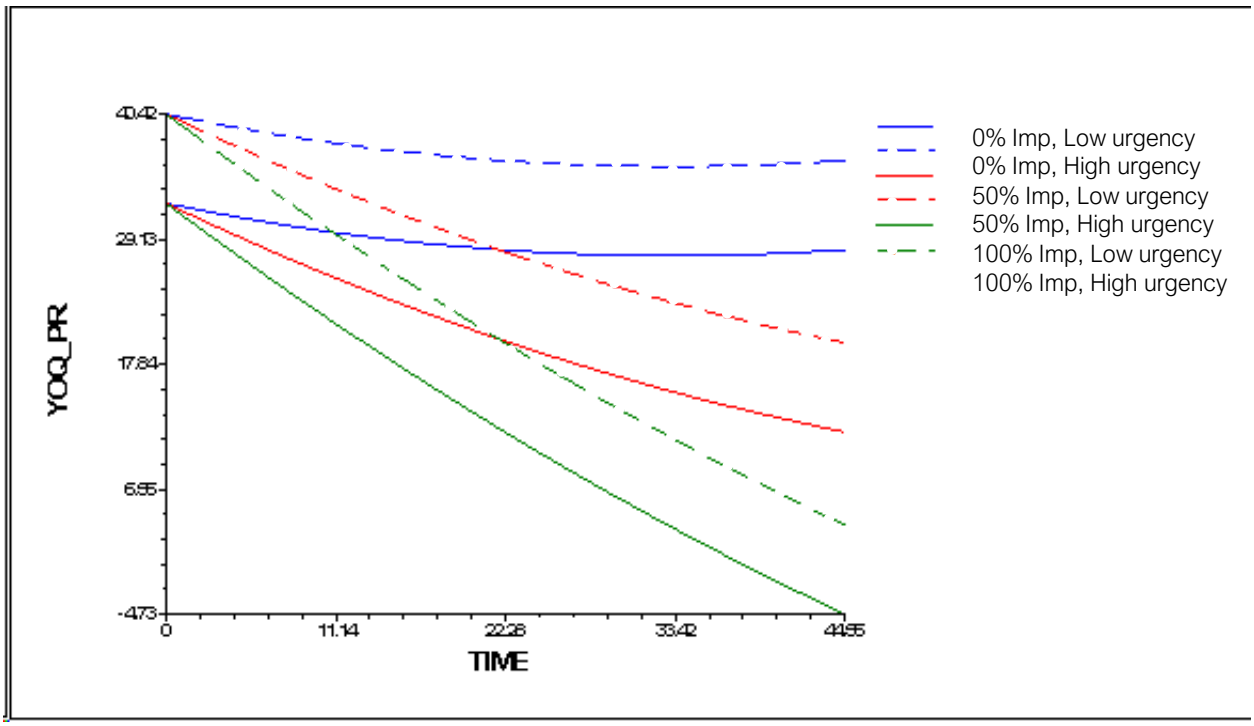



Figure 7. Impact of intervention dosage ("x% Imp") and patient urgency at intake (high vs. low) on parent-reported outcome scores over time. Intervention dosage has been split into three levels to illustrate level of implementation: 0 = no MFS implementation, 50 = partial MFS implementation, or intervention dose, and 100 = full MFS implementation and dose. Intervention dosage has been further separated into high and low urgency at intake scores.

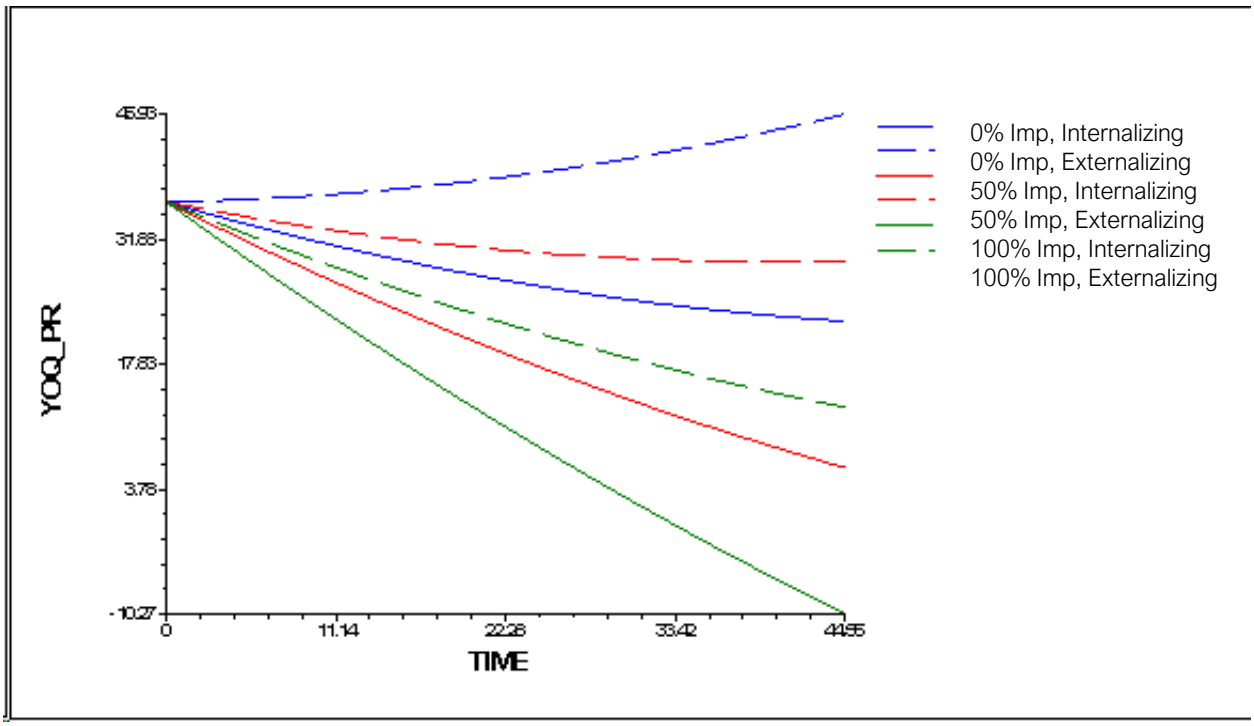



**Figure 8.** Impact of intervention dosage ("x% Imp") and diagnostic profile on parent-reported outcome scores over time. Implementation Index has been broken down into different levels that reflect intervention dosage: 0 = none, 50 = partial implementation, 100 = full implementation of the MFS and full intervention dose. These levels have then been further separated into internalizing and externalizing diagnostic types.

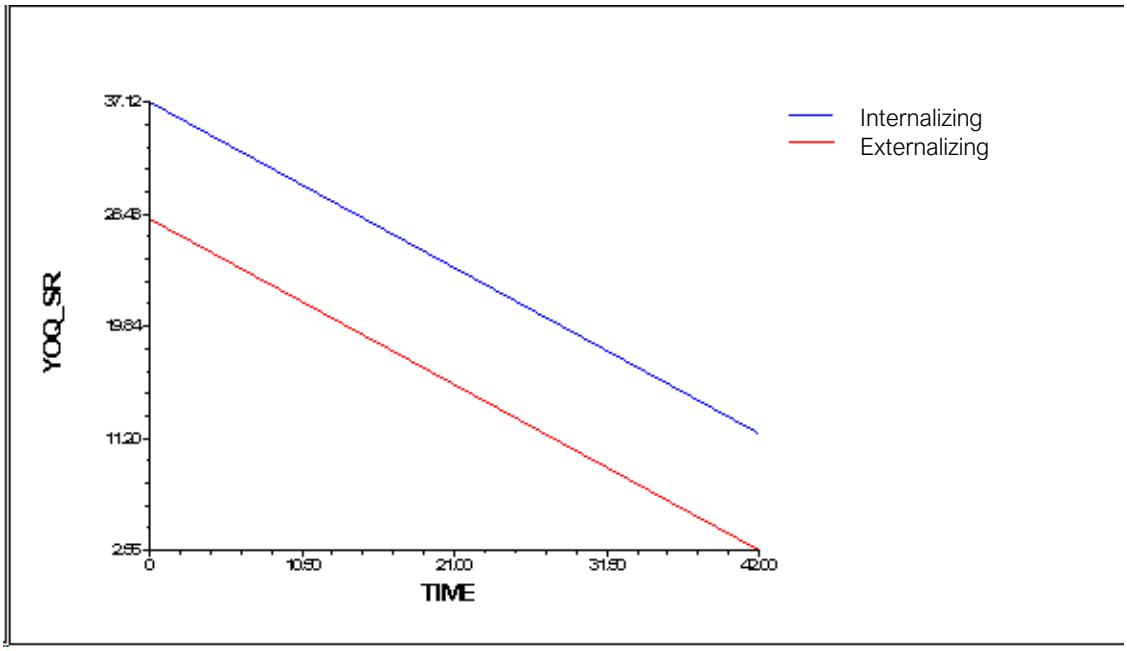



**Figure 9.** Change in outcome scores over time separated into diagnostic profile. One average, externalizing patients reported fewer symptoms than internalizing patients. Change over time was not significantly related to diagnosis type.

Table 1. Description, Explanation, and Interpretation of Implementation Index Examples

| Implementation Index | Examples of how the Index is obtained | Interpretation |
|---|---|---|
| 100 | A parent completed a questionnaire at every session and the clinician viewed every resulting feedback report | Ideal implementation: this patient received 100% of the intended intervention |
| 50 | A patient completed questionnaires at half of the sessions and the clinician viewed every resulting feedback report<br><br>OR<br><br>A patient completed a questionnaire at every session and the clinician only viewed half of the resulting feedback reports<br><br>OR<br><br>A patient completed 9 questionnaires at 10 sessions (90%) and the clinician viewed half of the resulting feedback reports | Partial implementation: this patient received 50% of the intended intervention |
| 0 | A parent did not complete a single questionnaire at any session<br><br>OR<br><br>The clinician did not view any feedback reports from questionnaires completed by a parent at every single session | Complete implementation failure: the patient received none of the intended intervention |

*Note.* Table and examples are adapted from Bickman et al., 2016. Examples provided are not exhaustive for obtaining an Implementation Index of 0, 50, 100.

Table 2. Participant Characteristics and Differences Across Waves

| Characteristic | Wave A (*n* = 279) M (SD) | Wave B (*n* = 259) M (SD) |
|---|---|---|
| Age at intake | 9.75 (3.74) | 10.01 (4.05) |
| Urgency at intake | 3.86 (1.04)* | 4.16 (1.31)* |
| Per treatment episode | | |
| Sessions | 8.32 (6.57) | 8.77 (6.00) |
| Weeks | 13.35 (8.30) | 14.79 (8.87) |
| Days | 93.44 (58.03) | 103.52 (61.95) |
| | *n* (%) | *n* (%) |
| Gender | | |
| Male | 157 (56.27) | 150 (57.92) |
| Female | 122 (43.73) | 108 (41.70) |
| Other | 0 | 1 (<1) |
| Ethnicity | | |
| Latinx/Hispanic | 139 (49.82) | 138 (53.28) |
| White/Non-Hispanic | 119 (42.65) | 89 (34.36) |
| African-American | 16 (5.73) | 26 (10.04) |
| Asian/Pacific-Islander | 3 (1.08) | 0 |
| Other | 1 (<1) | 4 (1.54) |
| History of trauma | | |
| Yes (With related diagnosis) | 59 (21.15) | 73 (28.19) |
| No | 220 (78.85) | 186 (71.82) |
| Diagnosis (Primary) | | |
| Internalizing | 159 (56.99) | 137 (52.90) |
| Externalizing | 118 (42.29) | 118 (45.56) |
| Other | 2 (<1) | 4 (1.54) |
| Comorbidity (% of patients with >1 diagnosis) | 52.90% | 62.01% |
| Insurance type | | |
| Subsidized | 239 (85.66) | 214 (82.60) |
| Private | 40 (14.34) | 45 (17.40) |

*$p < .01$

Table 3. Clinician Characteristics and Wave Assignment

| Characteristic | Wave A ($n$ = 34) M (SD) | Wave B ($n$ = 36) M (SD) |
|---|---|---|
| Caseload in study | 8.21 (6.39) | 7.26 (5.24) |
| Urgency per caseload | 3.86 (0.57)* | 4.17 (0.77)* |
| | *n* (%) | *n* (%) |
| Gender | | |
| Male | 2 (5.88) | 3 (8.82) |
| Female | 31 (91.18) | 31 (91.18) |
| Ethnicity | | |
| Latinx/Hispanic | 4 (11.76) | 7 (19.45) |
| White/Non-Hispanic | 28 (82.35) | 24 (66.67) |
| African-American | 1 (2.94) | 1 (2.78) |
| Asian/Pacific-Islander | 1 (2.94) | 1 (2.78) |
| Other | 0 | 3 (8.34) |
| Spanish-speaking | 9 (26.47) | 14 (38.89) |
| Supervisory status | | |
| Supervisors | 18 (52.94) | 16 (44.45) |
| Supervisees | 16 (47.06) | 20 (55.56) |
| Discipline | | |
| Social work | 15 (44.12) | 13 (36.12) |
| Counseling | 7 (20.59) | 7 (19.45) |
| School psychology | 7 (20.59) | 12 (33.34) |
| Clinical psychology | 5 (14.71) | 4 (11.12) |

*$p$ < .001. *Note.* Twenty-one clinicians were both in Wave A and Wave

Table 4. Hypothesized Growth Curve Models for Treatment Outcome Data

| | Parent-report | | | Youth-report | | |
|---|---|---|---|---|---|---|
| Fixed effects | B | SE | t | B | SE | t |
| Intercept | 32.958** | 1.080 | 30.518 | 32.225** | 1.654 | 19.483 |
| Urgency at intake | 2.139* | 0.845 | 2.531 | 0.611 | 1.514 | 0.404 |
| Interaction | 0.018 | 0.027 | 0.659 | -0.006 | 0.055 | -0.106 |
| For Time slope | | | | | | |
| Intercept | -0.387* | 0.186 | -2.081 | -0.080 | 0.290 | -0.277 |
| MFS dosage | -0.006* | 0.003 | -2.047 | -0.009 | 0.005 | -1.812 |
| Urgency at intake | -0.091 | 0.065 | -1.405 | -0.081 | 0.100 | -0.180 |
| Interaction | 0.001 | 0.005 | 0.363 | 0.005 | 0.004 | 1.362 |

| | Variance components | | | | | |
|---|---|---|---|---|---|---|
| Random Effects | Estimate | SD | $\chi^2$ | Estimate | SD | $\chi^2$ |
| $r_0$ | 252.634** | 15.895 | 8233.533 | 260.549** | 16.142 | 3482.494 |
| $r_1$ | 0.658** | 0.811 | 709.558 | 0.692 | 0.832 | 379.585 |

*$p < .05$ **$p < .001$

Table 5. Growth Curve Models with Hypothesized and Exploratory Predictors

| | Parent-report | | | Youth-report | | |
|---|---|---|---|---|---|---|
| Fixed effects | B | SE | t | B | SE | t |
| Intercept | 26.205*** | 2.246 | 11.669 | 36.995*** | 1.798 | 20.577 |
| Urgency at intake | 2.502* | 0.833 | 3.003 | 1.079 | 1.375 | 0.785 |
| Alert (NOT) | 9.234*** | 2.238 | 4.126 | | | |
| Diagnosis | 3.548 | 2.282 | 1.555 | -9.121* | 4.284 | -2.129 |
| MFSxURG | 0.011 | 0.026 | 0.422 | 0.013 | 0.050 | 0.260 |
| | | | | | | |
| For Time slope | | | | | | |
| Intercept | -0.562** | 0.183 | -3.073 | -0.313 | 0.306 | -1.022 |
| Diagnosis | 0.387* | 0.169 | 2.294 | | | |
| MFS dosage | -0.006 | 0.003 | -1.904 | -0.006 | 0.005 | -1.130 |
| MFSxURG | -0.000 | 0.003 | -0.094 | 0.004 | 0.004 | 0.986 |
| Urgency at intake | -0.120 | 0.066 | -1.831 | -0.044 | 0.109 | -0.402 |
| | | | | | | |
| For $Time^2$ slope | | | | | | |
| Intercept | 0.011* | 0.005 | 2.350 | | | |

| | Variance components | | | | | |
|---|---|---|---|---|---|---|
| Random Effects | Estimate | SD | $\chi^2$ | Estimate | SD | $\chi^2$ |
| $r_0$ | 228.325*** | 15.110 | 7108.956 | 252.301*** | 15.884 | 1241.274 |
| $r_1$ | 0.783*** | 0.613 | 669.555 | 0.688*** | 0.829 | 369.815 |

*$p$ < .05 **$p$ < .01 ***$p$ < .001

Table 6. Final Growth Curve Models for Treatment Outcome Data

| | Parent-report | | | Youth-report | | |
|---|---|---|---|---|---|---|
| Fixed effects | B | SE | t | B | SE | t |
| Intercept | 29.302*** | 1.308 | 22.408 | 37.119*** | 1.853 | 20.027 |
| Urgency at intake | 1.954* | 0.835 | 2.340 | | | |
| Alert (NOT) | 10.161*** | 2.248 | 4.520 | | | |
| Diagnosis | | | | -8.999** | 3.126 | -2.879 |
| For Time slope | | | | | | |
| Intercept | -0.505** | 0.185 | -2.696 | -0.609*** | 0.122 | -5.006 |
| Diagnosis | 0.360* | 0.178 | 2.023 | | | |
| MFS dosage | -0.006* | 0.003 | -2.114 | | | |
| For $Time^2$ slope | | | | | | |
| Intercept | 0.011* | 0.005 | 2.209 | | | |

| | Variance components | | | | | |
|---|---|---|---|---|---|---|
| Random Effects | Estimate | SD | $\chi^2$ | Estimate | SD | $\chi^2$ |
| $r_0$ | 233.496*** | 15.281 | 7171.945 | 247.508*** | 15.732 | 1250.041 |
| $r_1$ | 0.799*** | 0.640 | 698.050 | 0.693*** | 0.832 | 387.176 |

*$p < .05$ **$p < .01$ ***$p < .001$

# APPENDIX A

**Youth Outcome Questionnaire** **Name:** _____ **ID:** _____ **Date:** ___/___/_____
**Y-OQ®-30.2 SR TA English**

**PURPOSE:** The Y-OQ® 30.2 SR TA is designed to describe a wide range of troublesome situations, behaviors, and moods that are common to adolescents. You may discover that some of the items do not apply to your current situation. If so, please do not leave these items blank but mark the "Never or almost never" category.

When you begin to complete the Y-OQ® 30.2 SR TA, you will see that you can easily make yourself look as healthy or unhealthy as you wish. Please do not do that. If you are as accurate as possible it is more likely that you will be able to receive the help that you are seeking.

**DIRECTIONS:**

- Read each statement carefully.
- Decide how true this statement is during the past 7 days.
- Completely fill the circle that most accurately describes the past week.
- Fill in only one answer for each statement and erase unwanted marks clearly.

Please mark your answers like this:

O ● O

Not like this:

✖ ✔ ⊘

| | Never or Almost Never | Rarely | Sometimes | Frequently | Almost Always or Always |
|---|---|---|---|---|---|
| 1. I have headaches or feel dizzy. | O | O | O | O | O |
| 2. I don't participate in activities that used to be fun. | O | O | O | O | O |
| 3. I argue or speak rudely to others. | O | O | O | O | O |
| 4. I have a hard time finishing my assignments or I do them carelessly. | O | O | O | O | O |
| 5. My emotions are strong and change quickly. | O | O | O | O | O |
| 6. I have physical fights (hitting, kicking, biting, or scratching) with my family or others my age. | O | O | O | O | O |
| 7. I worry and can't get thoughts out of my mind. | O | O | O | O | O |
| 8. I steal or lie. | O | O | O | O | O |
| 9. I have a hard time sitting still (or I have too much energy). | O | O | O | O | O |
| 10. I use alcohol or drugs. | O | O | O | O | O |
| 11. I am tense and easily startled (jumpy). | O | O | O | O | O |
| 12. I am sad or unhappy. | O | O | O | O | O |
| 13. I have a hard time trusting friends, family members, or other adults. | O | O | O | O | O |
| 14. I think that others are trying to hurt me even when they are not. | O | O | O | O | O |
| 15. I have threatened to, or have run away from home. | O | O | O | O | O |
| 16. I physically fight with adults. | O | O | O | O | O |
| 17. My stomach hurts or I feel sick more than others my same age. | O | O | O | O | O |
| 18. I don't have friends or I don't keep friends very long. | O | O | O | O | O |
| 19. I think about suicide or feel I would be better off dead. | O | O | O | O | O |
| 20. I have nightmares, trouble getting to sleep, oversleeping, or waking up too early. | O | O | O | O | O |
| 21. I complain about or question rules, expectations, or responsibilities. | O | O | O | O | O |
| 22. I break rules, laws, or don't meet others' expectations on purpose. | O | O | O | O | O |
| 23. I feel irritated. | O | O | O | O | O |
| 24. I get angry enough to threaten others. | O | O | O | O | O |
| 25. I get into trouble when I'm bored. | O | O | O | O | O |
| 26. I destroy property on purpose. | O | O | O | O | O |
| 27. I have a hard time concentrating, thinking clearly, or sticking to tasks. | O | O | O | O | O |
| 28. I withdraw from my family and friends. | O | O | O | O | O |
| 29. I act without thinking and don't worry about what will happen. | O | O | O | O | O |
| 30. I feel like I don't have any friends or that no one likes me. | O | O | O | O | O |



For More Information Contact:

OQ Measures, LLC
P.O. Box 521047
Salt Lake City, UT 84152

Toll-Free USA: 1-888-MH-SCORE
(1-888-647-2673)

Phone: (801) 649-4392
Fax: (801) 747-6900
Email: INFO@OQMEASURES.COM
Website:
HTTP://WWW.OQMEASURES.COM

**Youth Outcome Questionnaire** **Name:** _____ **ID:** _____ **Date:** ___/___/_____
**Y-OQ®-30.2 SR TA English**

**PURPOSE:** The Y-OQ® 30.2 SR TA is designed to describe a wide range of troublesome situations, behaviors, and moods that are common to adolescents. You may discover that some of the items do not apply to your current situation. If so, please do not leave these items blank but mark the "Never or almost never" category.

When you begin to complete the Y-OQ® 30.2 SR TA, you will see that you can easily make yourself look as healthy or unhealthy as you wish. Please do not do that. If you are as accurate as possible it is more likely that you will be able to receive the help that you are seeking.

**DIRECTIONS:**

- Read each statement carefully.
- Decide how true this statement is during the past 7 days.
- Completely fill the circle that most accurately describes the past week.
- Fill in only one answer for each statement and erase unwanted marks clearly.

Please mark your answers like this:

○ ● ○

Not like this:

✖ ✔ ⊖

| | Strongly Disagree | Slightly Disagree | Neutral | Slightly Agree | Strongly Agree | Haven't Met Therapist Yet |
|---|---|---|---|---|---|---|
| 31. I feel like my therapist is on my side and tries to help me..... | ○ | ○ | ○ | ○ | ○ | ○ |
| 32. I look forward to meeting with my therapist. | ○ | ○ | ○ | ○ | ○ | ○ |
| 33. I don't feel I'm making much progress with my therapist..... | ○ | ○ | ○ | ○ | ○ | ○ |
| 34. I feel like my therapist knows how to help me. | ○ | ○ | ○ | ○ | ○ | ○ |
| 35. My therapist really listens to me..... | ○ | ○ | ○ | ○ | ○ | ○ |



For More Information Contact:

OQ Measures, LLC
P.O. Box 521047
Salt Lake City, UT 84152

Toll-Free USA: 1-888-MH-SCORE
(1-888-647-2673)

Phone: (801) 649-4392
Fax: (801) 747-6500
Email: INFO@OQMEASURES.COM
Website:
HTTP://WWW.OQMEASURES.COM

# APPENDIX B

**Youth Outcome Questionnaire** Name: _____ ID: _____ Date: ___/___/_____
**Y-OQ® 30.2 PR TA English Form**

**Youth Outcome Questionnaire** Name: _____ ID: _____ Date: ___/___/_____
**Y-OQ® 30.2 PR TA English Form**

**PURPOSE:** The Y-OQ® 30.2 PR TA is designed to describe a wide range of situations, behaviors, and moods that are common to children and adolescents. You may discover that some of the items do not apply to your current situation. If so, please do not leave these items blank but mark the "Never or almost never" category. When you begin to complete the Y-OQ® 30.2 PR TA you will see that you can easily make your child look as healthy or unhealthy as you wish. Please do not do that. If you are as accurate as possible it is more likely that you will be able to receive the help that you are seeking.

**DIRECTIONS:**

- Read each statement carefully.
- Decide how true this statement is during the past 7 days.
- Completely fill the circle that most accurately describes the past week.
- Fill in only one answer for each statement and erase unwanted marks clearly.

Please mark your answers like this:

○ ● ○

Not like this:

✖ ✔ ⊘



For More Information Contact:

OQ Measures, LLC
P.O. Box 521047
Salt Lake City, UT 84152

Toll-Free USA: 1-888-MH-SCORE
(1-888-647-2673)

Phone: (801) 649-4392
Fax: (801) 747-6900
Email: INFO@OQMEASURES.COM
Website:
HTTP://WWW.OQMEASURES.COM

| | Never or Almost Never | Rarely | Sometimes | Frequently | Almost Always or Always |
|---|---|---|---|---|---|
| 1. My child has headaches or feels dizzy. | O | O | O | O | O |
| 2. My child doesn't participate in activities that used to be fun..... | O | O | O | O | O |
| 3. My child argues or speaks rudely to others. | O | O | O | O | O |
| 4. My child has a hard time finishing assignments or does..... them carelessly. | O | O | O | O | O |
| 5. My child's emotions are strong and change quickly. | O | O | O | O | O |
| 6. My child has physical fights (hitting, kicking, biting, scratching) with family or others his/her age. | O | O | O | O | O |
| 7. My child worries and can't get thoughts out of his/her mind. | O | O | O | O | O |
| 8. My child steals or lies..... | O | O | O | O | O |
| 9. My child is has a hard time sitting still (or has too much energy). | O | O | O | O | O |
| 10. My child uses alcohol or drugs..... | O | O | O | O | O |
| 11. My child seems tense and easily startled (jumpy). | O | O | O | O | O |
| 12. My child is sad or unhappy..... | O | O | O | O | O |
| 13. My child has a hard time trusting friends, family members, or other adults. | O | O | O | O | O |
| 14. My child thinks that others are trying to hurt him/her even ..... when they're not. | O | O | O | O | O |
| 15. My child has threatened to, or has run away from home. | O | O | O | O | O |
| 16. My child physically fights with adults..... | O | O | O | O | O |
| 17. My child's stomach hurts or feels sick more than others his/her age. | O | O | O | O | O |
| 18. My child doesn't have friends or doesn't keep friends very long... | O | O | O | O | O |
| 19. My child thinks about suicide or feels s/he would be better off dead. | O | O | O | O | O |
| 20. My child has nightmares, trouble getting to sleep, oversleeping, or waking up too early. | O | O | O | O | O |
| 21. My child complains or questions rules, expectations, or responsibilities. | O | O | O | O | O |
| 22. My child breaks rules, laws, or doesn't meet others expectations on purpose | O | O | O | O | O |
| 23. My child feels irritated | O | O | O | O | O |
| 24. My child gets angry enough to threaten others..... | O | O | O | O | O |
| 25. My child gets into trouble when he/she is bored. | O | O | O | O | O |
| 26. My child destroys property on purpose..... | O | O | O | O | O |
| 27. My child has a hard time concentrating, thinking clearly, or sticking to tasks. | O | O | O | O | O |
| 28. My child withdraws from family and friends..... | O | O | O | O | O |
| 29. My child acts without thinking & doesn't worry about what will happen | O | O | O | O | O |
| 30. My child feels like s/he doesn't have any friends and that no one likes them | O | O | O | O | O |

**Youth Outcome Questionnaire** **Name:** _____ **ID:** _____ **Date:** ___/___/_____
**Y-OQ®-30.2 PR TA English Form**

**PURPOSE:** The Y-OQ® 30.2 PR TA is designed to describe a wide range of situations, behaviors, and moods that are common to children and adolescents. You may discover that some of the items do not apply to your current situation. If so, please do not leave these items blank but mark the "Never or almost never" category.

When you begin to complete the Y-OQ® 30.2 PR TA you will see that you can easily make your child look as healthy or unhealthy as you wish. Please do not do that. If you are as accurate as possible it is more likely that you will be able to receive the help that you are seeking.

**DIRECTIONS:**

- Read each statement carefully.
- Decide how true this statement is during the past 7 days.
- Completely fill the circle that most accurately describes the past week.
- Fill in only one answer for each statement and erase unwanted marks clearly.

Please mark your answers like this:

O ● O

Not like this:

✖ ✔ ⊖

| | Strongly Disagree | Slightly Disagree | Neutral | Slightly Agree | Strongly Agree | Haven't Met Therapist Yet |
|---|---|---|---|---|---|---|
| 31. I look forward to meeting my child's therapist..... | O | O | O | O | O | O |
| 32. I don't feel like my child is making much progress with his/her therapist | O | O | O | O | O | O |
| 33. I feel like my child's therapist knows how to help my child..... | O | O | O | O | O | O |
| 34. My child's therapist really listens to me..... | O | O | O | O | O | O |



For More Information Contact:

OQ Measures, LLC
P.O. Box 521047
Salt Lake City, UT 84152

Toll-Free USA: 1-888-MH-SCORE
(1-888-647-2673)

Phone: (801) 649-4392
Fax: (801) 747-6900
Email: INFO@OQMEASURES.COM
Website:
HTTP://WWW.OQMEASURES.COM

## References


Aiken, L. S., & West, S. G. (1991). Multiple regression: Testing and interpreting interactions. Newbury Park, CA: Sage.

Amble, I., Gude, T., Stubdal, S., Andersen, B. J., & Wampold, B. E. (2015). The effect of implementing the Outcome Questionnaire-45.2 feedback system in Norway: A multisite randomized clinical trial in a naturalistic setting. *Psychotherapy Research, 25*(6), 669-677. doi:10.1080/10503307.2014.928756

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). doi:10.1176/appi.books.9780890423349.

American Psychological Association. (2002). Ethical Principles of Psychologists and Code of Conduct. *American Psychologist, 57*(12), 1060-1073.

American Psychological Association (2006). Evidence-based practice in psychology. *American Psychologist*, *61*(4), 271-285. doi:10.1037/0003-066X.61.4.271

Bar-Kalifa, E., Atzil-Slonim, D., Rafaeli, E., Peri, T., Rubel, J., & Lutz, W. (2016). Therapist–client agreement in assessments of clients' functioning. *Journal of Consulting and Clinical Psychology*, *84*(12), 1127.

Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., . . . McGrath, G. (2001). Service profiling and outcomes benchmarking using the CORE-OM: Towards practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology, 69*, 184-196. doi:10.1037/0022-006X.69.2.184

Batty, M. J., Moldavsky, M., Foroushani, P. S., Pass, S., Marriott, M., Sayal, K., & Hollis, C. (2013). Implementing routine outcome measures in child and adolescent mental health services: from present to future practice. *Child & Adolescent Mental Health*, *18*(2), 82-87.

Bearman, S. K., & Weisz, J. R. (2015). Review: Comprehensive treatments for youth comorbidity—Evidence-guided approaches to a complicated problem. *Child and Adolescent Mental Health*, *20*(3), 131-141. doi:10.1111/camh.12092

Beck, J. G., Castonguay, L. G., Chronis-Tuscano, A., Klonsky, E. D., McGinn, L. K., & Youngstrom, E. A. (2014). Principles for training in evidence-based psychology: Recommendations for the graduate curricula in clinical psychology. *Clinical Psychology: Science and Practice, 21*(4), 410-424. doi:10.1111/cpsp.12079

Becker-Haimes, E. M., Okamura, K. H., Baldwin, C. D., Wahesh, E., Schmidt, C., & Beidas, R. S. (2018). Understanding the landscape of behavioral health pre-service training to inform evidence-based intervention implementation. *Psychiatric Services, 70*(1), 1-3.

Beidas, R. S., Edmunds, J. M., Marcus, S. C., & Kendall, P. C. (2012). Training and consultation to promote implementation of an empirically supported treatment: A randomized trial. *Psychiatric Services*, *63*(7), 660-665.

Bergman, H., Kornør, H., Nikolakopoulou, A., Hanssen-Bauer, K., Soares-Weiser, K., Tollefsen, T. K., & Bjørndal, A. (2018). Client feedback in psychological therapy for children and adolescents with mental health problems. *Cochrane Database of*

*Systematic Reviews*, *8*, CD011729-CD011729. doi:10.1002/14651858.CD011729.pub2

Bickman, L. (2008). Why don't we have effective mental health services? *Administration and Policy in Mental Health and Mental Health Services Research*, *35*(6), 437-439. doi:10.1007/s10488-008-0192-9

Bickman, L., Douglas, S. R., De Andrade, A. V., Tomlinson, M., Gleacher, A., Olin, S., & Hoagwood, K. (2016). Implementing a measurement feedback system: A tale of two sites. *Administration and Policy in Mental Health and Mental Health Services Research*, *43*(3), 410-425. doi:10.1007/s10488-015-0647-8

Bickman, L., Kelley, S. D., Breda, C., de Andrade, A. R., & Riemer, M. (2011). Effects of routine feedback to clinicians on mental health outcomes of youths: Results of a randomized trial. *Psychiatric Services*,*62*(12), 1423-1429. doi:10.1176/appi.ps.0020 52011

Bickman, L., & Noser, K. (1999). Meeting the challenges in the delivery of child and adolescent mental health services in the next millennium: The continuous quality improvement approach. *Applied & Preventive Psychology*, *8*(4), 247-255. doi:10.1016/S0962-1849(05)80039-3

Bickman, L., Rosof-Williams, J., Salzer, M. S., Summerfelt, W. T., Noser, K., Wilson, S. J., & Karver, M. S. (2000). What information do clinicians value for monitoring adolescent client progress and outcomes? *Professional Psychology: Research and Practice*, *31*(1), 70-74. doi:10.1037/0735-7028.31.1.70

Birmaher, B., & Brent, D. (2007). Practice parameter for the assessment and treatment of children and adolescents with depressive disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, *46*(11), 1503-1526.

Borntrager, C., Chorpita, B. F., Higa-McMillan, C. K., Daleiden, E. L., & Starace, N. (2013). Usual care for trauma-exposed youth: Are clinician-reported therapy techniques evidence-based? *Children and Youth Services Review, 35*(1), 133-141. 10.1016/j.childyouth.2012.09.018

Borntrager, C., & Lyon, A. R. (2015). Client progress monitoring and feedback in school-based mental health. *Cognitive and Behavioral Practice*, *22*(1), 74-86. doi:10.1016/j.cbpra. 2014.03.007

Boswell, J. F., Kraus, D. R., Miller, S. D., & Lambert, M. J. (2015). Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. *Psychotherapy Research*, *25*(1), 6-19. doi:10.1080/10503307.2013. 817696

Brattland, H., Koksvik, J. M., Burkeland, O., Gråwe, R. W., Klöckner, C., Linaker, O. M., … Iversen, V. C. (2018). The effects of routine outcome monitoring (ROM) on therapy outcomes in the course of an implementation process: A randomized clinical trial. *Journal of Counseling Psychology*, *65*(5), 641–652.

Brown, G. S., & Jones, E. R. (2005). Implementation of a feedback system in a managed care environment: What are patients teaching us? *Journal of Clinical Psychology, 61*(2), 187-198. doi:10.1002/jclp.20110

Burlingame, G. M., Dunn, T., Cox, J., Wells, G., Lambert, M. J., & Brown, G. S. (2004). *Administration and scoring manual for the Youth Outcome Questionnaire-30 (Y-OQ-30).* Salt Lake City, UT: Outcome Questionnaire Measures.

Burlingame, G. M., Mosier, J. I., Wells, G. M., Atkin, Q. G., Lambert, M. J., Whoolery, M., & Latkowski, M. (2001). Tracking the influence of mental health treatment: The development of the Youth Outcome Questionnaire. *Clinical Psychology & Psychotherapy*, *8*(5), 361-379.

Cannon, J. N., Warren, J. S., Nelson, P. L., & Burlingame, G. M. (2010). Change trajectories for the Youth Outcome Questionnaire Self-Report: Identifying youth at risk for treatment failure. *Journal of Clinical Child and Adolescent Psychology*, *39*(3), 289-301. doi:10.1080/15374411003691727

Carlier, I. E., Meuldijk, D., Van Vliet, I. M., Van Fenema, E., Van der Wee, N. A., & Zitman, F. G. (2012). Routine outcome monitoring and feedback on physical or mental health status: Evidence and theory. *Journal of Evaluation in Clinical Practice*, *18*(1), 104-110. doi:10.1111/j.1365-2753.2010.01543.x

Centers for Disease Control and Prevention (2016). Mental health and substance use disorder expenditures, by type of expenditure: United States, selected years 1986-2014. Retrieved from https://www.cdc.gov/nchs/hus/contents2016.htm#fig24

Chorpita, B. F., Daleiden, E. L., Park, A. L., Ward, A. M., Levy, M. C., Cromley, T., . . . Krull, J. L. (2017). Child STEPs in California: A cluster randomized effectiveness trial comparing modular treatment with community implemented treatment for

youth with anxiety, depression, conduct problems, or traumatic stress. *Journal of Consulting and Clinical Psychology*, *85*(1), 13-25.

Chorpita, B. F., Daleiden, E. L., & Weisz, J. R. (2005). Identifying and selecting the common elements of evidence based interventions: A distillation and matching model. *Mental health services research*, *7*(1), 5-20.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Connolly, S. D., & Bernstein, G. A. (2007). Practice parameter for the assessment and treatment of children and adolescents with anxiety disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, *46*(2), 267-283.

Connolly-Gibbons, M. B., Kurtz, J. E., Thompson, D. L., Mack, R. A., Lee, J. K., Rothbard, A., & ... Crits-Christoph, P. (2015). The effectiveness of clinician feedback in the treatment of depression in the community mental health system. *Journal of Consulting and Clinical Psychology*, *83*(4), 748-759. doi:10.1037/a0039302

Cooper, M., Stewart, D., Sparks, J., & Bunting, L. (2013). School-based counseling using systematic feedback: A cohort study evaluating outcomes and predictors of change. *Psychotherapy Research*, *23*(4), 474-488. doi:10.1080/10503307.2012.735777

Costello, E. J., He, J., Sampson, N. A., Kessler, R. C., & Merikangas, K. R. (2014). Services for adolescents with psychiatric disorders: 12-month data from the National Comorbidity Survey—Adolescent. *Psychiatric Services*, *65*(3), 359-366.

Costello, E. J., Mustillo, S., Erkanli, A., Keeler, G., & Angold, A. (2003). Prevalence and development of psychiatric disorders in childhood and adolescence. *Archives of General Psychiatry*, *60*(8), 837– 844.

Crits-Christoph, P., Ring-Kurtz, S., McClure, B., Temes, C., Kulaga, A., Gallop, R., & ... Rotrosen, J. (2010). A randomized controlled study of a web-based performance improvement system for substance abuse treatment providers. *Journal of Substance Abuse Treatment*, *38*(3), 251-262. doi:10.1016/j.jsat.2010.01.001

Davidson, K. M., Rankin, M. L., Begley, A., Lloyd, S., Barry, S. E., McSkimming, P., & ... Walker, A. (2017). Assessing patient progress in psychological therapy through feedback in supervision: The MeMOS (Measuring and Monitoring clinical Outcomes in Supervision) randomized controlled trial. *Behavioural and Cognitive Psychotherapy*, *45*(3), 209-224. doi:10.1017/S1352465817000029

De Haan, A. M., Boon, A. E., De Jong, J. T. V. M., Hoeve, M., & Vermeiren, R. R. J. M. (2013). A meta-analytic review on treatment dropout in child and adolescent outpatient mental health care. *Clinical Psychology Review, 33*(5), 698–711.

De Jong, K., Timman, R., Hakkaart-Van Roijen, L., Vermeulen, P., Kooiman, K., Passchier, J., & Van Busschbach, J. (2014). The effect of outcome monitoring feedback to clinicians and patients in short and long-term psychotherapy: A randomized controlled trial. *Psychotherapy Research*, *24*(6), 629-639. doi:10. 1080/10503307.2013.871079

De Jong, K., van Sluis, P., Nugter, M. A., Heiser, W. J., & Spinhoven, P. (2012). Understanding the differential impact of outcome monitoring: Therapist variables

that moderate feedback effects in a randomized clinical trial. *Psychotherapy Research*, *22*(4), 464-474. doi:10.1080/10503307.2012.673023

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin*, *131*(4), 483.

Delgadillo, J., de Jong, K., Lucock, M., Lutz, W., Rubel, J., Gilbody, S., . . . McMillan, D. (2018). Feedback-informed treatment versus usual psychological treatment for depression and anxiety: A multisite, open-label, cluster randomised controlled trial. *The Lancet Psychiatry, 5*(7), 564-572. doi:10.1016/S2215-0366(18)30162-7

Delgadillo, J., Overend, K., Lucock, M., Groom, M., Kirby, N., McMillan, D., & ... de Jong, K. (2017). Improving the efficiency of psychological treatment using outcome feedback technology. *Behaviour Research and Therapy, 99*, 89-97. doi:10.1016/j.brat.2017.09.011

Douglas, S. R., Jonghyuk, B., de Andrade, A. R. V., Tomlinson, M. M., Hargraves, R. P., & Bickman, L. (2015). Feedback mechanisms of change: How problem alerts reported by youth clients and their caregivers impact clinician-reported session content. *Psychotherapy Research*, *25*(6), 678-693.

Edbrooke-Childs, J., Wolpert, M., & Deighton, J. (2016). Introduction to the special section on implementing feedback from outcome measures in child and adult mental health services. *Administration and Policy in Mental Health and Mental Health Services Research, 43*(3), 277-278. doi:10.1007/s10488-016-0726-5

Gabbard, G. O., Lazar, S. G., Hornberger, J., & Spiegel, D. (1997). The economic impact of psychotherapy: A review. *The American Journal of Psychiatry*, *154*(2), 147.

Garland, A. F., Brookman-Frazee, L., Hurlburt, M., Acccurso, E., Zoffness, R., Haine-Schlagel, R., & Ganger, W. (2010). Mental health care for children with disruptive behavior problems: A view inside therapists' offices. *Psychiatric Services, 61,* 788–795.

Garland, A. F., Lau, A. S., Yeh, M., McCabe, K. M., Hough, R. L., & Landsverk, J. A. (2005). Racial and ethnic differences in utilization of mental health services among high-risk youths. *American Journal of Psychiatry*, *162*(7), 1336-1343.

Garland, A. F., Lebensohn-Chialvo, F., Hall, K. G., & Cameron, E. R. (2017). Capitalizing on scientific advances to improve access to and quality of children's mental health care. *Behavioral Sciences & the Law*, *35*(4), 337-352.

Gleacher, A. A., Olin, S. S., Nadeem, E., Pollock, M., Ringle, V., Bickman, L., & ... Hoagwood, K. (2016). Implementing a measurement feedback system in community mental health clinics: A case study of multilevel barriers and facilitators. *Administration and Policy in Mental Health and Mental Health Services Research*, *43*(3), 426-440. doi:10.1007/s10488-015-0642-0

Goldberg, S. B., Babins-Wagner, R., Rousmaniere, T., Berzins, S., Hoyt, W. T., Whipple, J. L., & ... Wampold, B. E. (2016). Creating a climate for therapist improvement: A case study of an agency focused on outcomes and deliberate practice. *Psychotherapy*, *53*(3), 367-375. doi:10.1037/pst0000060

Gondek, D., Edbrooke-Childs, J., Fink, E., Deighton, J., & Wolpert, M. (2016). Feedback from outcome measures and treatment effectiveness, treatment efficiency, and collaborative practice: A systematic review. *Administration and Policy in Mental Health and Mental Health Services Research, 43*(3), 325-343. doi:10.1007/s10488-015-0710-5

Grossl, A. B., Reese, R. J., Norsworthy, L. A., & Hopkins, N. B. (2014). Client feedback data in supervision: Effects on supervision and outcome. *Training and Education in Professional Psychology, 8*(3), 182-188. doi:10.1037/tep0000025

Hall, C. L., Taylor, J., Moldavsky, M., Marriott, M., Pass, S., Newell, K., & ... Hollis, C. (2014). A qualitative process evaluation of electronic session-by-session outcome measurement in child and adolescent mental health services. *BMC Psychiatry, 14.*

Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. W., Shimokawa, K., & Sutton, S. W. (2005). A lab test and algorithms for identifying clients at risk for treatment failure. *Journal of Clinical Psychology*, *61*(2), 155-163.

Harkin, B., Webb, T. L., Chang, B. I., Prestwich, A., Conner, M., Kellar, I., & ... Sheeran, P. (2016). Does monitoring goal progress promote goal attainment? A meta-analysis of the experimental evidence. *Psychological Bulletin*, *142*(2), 198-229. doi:10.1037/bul0000025

Hatfield, D., McCullough, L., Frantz, S. B., & Krieger, K. (2010). Do we know when our clients get worse? An investigation of therapists' ability to detect negative client change. *Clinical Psychology & Psychotherapy*, *17*(1), 25-32.

Hawley, K. M., & Weisz, J. R. (2003). Child, parent and therapist (dis) agreement on target problems in outpatient therapy: The therapist's dilemma and its implications. *Journal of Consulting and Clinical Psychology*, *71*(1), 62.

Holder, S. M., Rogers, K., Peterson, E., & Ochonma, C. (2017). Mental health visits: Examining socio-demographic and diagnosis trends in the emergency department by the pediatric population. *Child Psychiatry & Human Development*, *48*(6), 993-1000.

Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy*, *48*(1), 9-22.

Ionita, G., Fitzpatrick, M., Tomaro, J., Chen, V. V., & Overington, L. (2016). Challenges of using progress monitoring measures: Insights from practicing clinicians. *Journal of Counseling Psychology*, *63*(2), 173.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*, 12-19.

Janse, P. D., De Jong, K., Van Dijk, M. K., Hutschemaekers, G., & Verbraak, M. (2017). Improving the efficiency of cognitive behavioural therapy by using formal client feedback, *Psychotherapy Research*, *27*(5), 525-538. doi:10.1080/10503307.2016.1152408

Jensen-Doss, A., Haimes, E. B., Smith, A. M., Lyon, A. R., Lewis, C. C., Stanick, C. F., & Hawley, K. M. (2016). Monitoring treatment progress and providing feedback is viewed favorably but rarely used in practice. *Administration and Policy in*

*Mental Health and Mental Health Services Research*. doi:10.1007/s10488-016-0763-0

Jensen-Doss, A., Smith, A. M., Becker-Haimes, E. M., Mora Ringle, V., Walsh, L. M., Nanda, M., & ... Lyon, A. R. (2017). Individualized progress measures are more acceptable to clinicians than standardized measures: Results of a national survey. *Administration and Policy in Mental Health and Mental Health Services Research*, *45*(3), 392-403. doi:10.1007/s10488-017-0833-y

Kalb, L. G., Stapp, E. K., Ballard, E. D., Holingue, C., Keefer, A., & Riley, A. (2019). Trends in psychiatric emergency department visits among youth and young adults in the US. *Pediatrics*, *143*(4), e20182192.

Kapke, T. L., & Gerdes, A. C. (2016). Latino family participation in youth mental health services: Treatment retention, engagement, and response. *Clinical child and family psychology review*, *19*(4), 329-351.

Kelley, S. D., & Bickman, L. (2009). Beyond outcomes monitoring: Measurement feedback systems in child and adolescent clinical practice. *Current Opinion in Psychiatry*, *22*(4), 363-368. doi:10.1097/YCO.0b013e32832c9162

Kelley, S. D., de Andrade, A. R. V., Bickman, L., & Robin, A. V. (2012). The Session Report Form (SRF): Are clinicians addressing concerns reported by youth and caregivers. *Administration and Policy in Mental Health and Mental Health Services Research*, *39*(1-2), 133-145.

Kendrick, T., El-Gohary, M., Stuart, B., Gilbody, S., Churchill, R., Aiken, L., & ... Moore, M. (2016). Routine use of patient reported outcome measures (PROMs)

for improving treatment of common mental health disorders in adults. *The Cochrane Database of Systematic Reviews*, *7*CD011119. doi:10.1002/14651858.CD011119

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254.

Knaup, C., Koesters, M., Schoefer, D., Becker, T., & Puschner, B. (2009). Effect of feedback of treatment outcome in specialist mental healthcare: Meta-analysis. *The British Journal of Psychiatry*, *195*(1), 15-22.

Kotte, A., Hill, K. A., Mah, A. C., Korathu-larson, P., Au, J. R., Izmirian, S., . . . Higa-McMillan, C. (2016). Facilitators and barriers of implementing a measurement feedback system in public youth mental health. *Administration and Policy in Mental Health and Mental Health Services Research*, *43*(6), 861-878. doi: 10.1007/s10488-016-0729-2

Lambert, M. J. (2015). Progress feedback and the OQ-system: The past and the future. *Psychotherapy*, *52*(4), 381-390. doi:10.1037/pst0000027

Lambert, M. J., Burlingame, G. M., Umphress, V., Hansen, N. B., Vermeersch, D. A., Clouse, G. C., & Yanchar, S. C. (1996). The reliability and validity of the outcome questionnaire. *Clinical Psychology & Psychotherapy*, *3*, 249-258. doi:10.1002/(SICI)1099-0879(199612)3:4<249::AID-CPP106>3.0.CO;2-S

Lambert, M., Hansen, N., & Harmon, S. (2010). Outcome questionnaire system (The OQ System): Development and practical applications in healthcare settings. In M.

Barkham, G.E. Hardy, & J. Mellor-Clark (Eds), *Developing and delivering practice-based evidence: A guide for the psychological therapies,* (pp. 141–154). Chichester: John Wiley & Sons.

Lambert, M. J., Whipple, J. L., Bishop, M. J., Vermeersch, D. A., Gray, G. V., & Finch, A. E. (2002a). Comparison of empirically-derived and rationally-derived methods for identifying patients at risk for treatment failure. *Clinical Psychology & Psychotherapy*, *9*(3), 149-164. doi:10.1002/cpp.333

Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen, S. L., & Smart, D. W. (2003). Is it time for clinicians to routinely track patient outcome? A meta-analysis. *Clinical Psychology: Science and Practice*, *10*(3), 288-301. doi:10.1093/clipsy/bpg025

Lambert, M. J., Whipple, J. L., & Kleinstäuber, M. (2018). Collecting and delivering progress feedback: A meta-analysis of routine outcome monitoring. *Psychotherapy (Chicago, Ill.), 55*(4), 520-537. doi:10.1037/pst0000167

Lambert, M. J., Whipple, J. L., Vermeersch, D. A., Smart, D. W., Hawkins, E. J., Nielsen, S. L., & Goates, M. (2002b). Enhancing psychotherapy outcomes via providing feedback on client progress: A replication. *Clinical Psychology & Psychotherapy*, *9*(2), 91-103. doi:10.1002/cpp.324

Lambert, M. J., & Shimokawa, K. (2011). Collection of client feedback. *Psychotherapy, 48*(1), 72.

Leslie, D. L., & Rosenheck, R. (1999). Shifting to outpatient care? Mental health care use and cost under private insurance. *American Journal of Psychiatry*, *156*(8), 1250-1257.

Lucock, M., Halstead, J., Leach, C., Barkham, M., Tucker, S., Randal, C., & ... Saxon, D. (2015). A mixed-method investigation of patient monitoring and enhanced feedback in routine practice: Barriers and facilitators. *Psychotherapy Research*, *25*(6), 633-646. doi:10.1080/10503307.2015.1051163

Lutz, W., Rubel, J., Schiefele, A., Zimmermann, D., Böhnke, J. R., & Wittmann, W. W. (2015). Feedback and therapist effects in the context of treatment outcome and treatment length. *Psychotherapy Research*, *25*(6), 647-660. doi:10.1080/10503307.2015.1053553

Lyon, A. R., Lewis, C. C., Boyd, M. R., Hendrix, E., & Liu, F. (2016). Capabilities and characteristics of digital measurement feedback systems: Results from a comprehensive review. *Administration and Policy in Mental Health and Mental Health Services Research*, *43*(3), 441-466. doi:10.1007/s10488-016-0719

Macdonald, J., & Mellor-Clark, J. (2015). Correcting psychotherapists' blindsidedness: Formal feedback as a means of overcoming the natural limitations of therapists. *Clinical Psychology & Psychotherapy*, *22*(3), 249-257.

McClintock, A. S., Perlman, M. R., McCarrick, S. M., Anderson, T., & Himawan, L. (2017). Enhancing psychotherapy process with common factors feedback: A randomized, clinical trial. *Journal of Counseling Psychology*, *64*(3), 247-260. doi:10.1037/cou0000188

Miller, S. D., Duncan, B. L., Sorrell, R., & Brown, G. S. (2005). The Partners for Change Outcome Management System. *Journal of Clinical Psychology: In-Session, 61*, 199-208. doi:10.1002/jclp.20111

Mosier, J., Burlingame, G. M., Wells, M. G., Ferre, R., Latkowski, M., Johansen, J., ... & Walton, E. (2001). In-home, family-centered psychiatric treatment for high-risk children and youth. *Children's Services: Social Policy, Research, and Practice*, *4*(2), 51-68.

Ng, M. Y., & Weisz, J. R. (2016). Annual research review: Building a science of personalized intervention for youth mental health. *Journal of Child Psychology & Psychiatry*, *57*(3), 216-236. doi:10.1111/jcpp.12470

Norcross, J. C., & Wampold, B. E. (2011). What works for whom: Tailoring psychotherapy to the person. *Journal of Clinical Psychology*, *67*(2), 127-132.

Norman, S., Dean, S., Hansford, L., & Ford, T. (2014). Clinical practitioner's attitudes towards the use of routine outcome monitoring within child and adolescent mental health services: A qualitative study of two child and adolescent mental health services. *Clinical Child Psychology and Psychiatry*, *19*(4), 576-595. doi:10.1177/1359104513492348

Pinner, D. H., & Kivlighan, D. M., III (2018). The ethical implications and utility of routine outcome monitoring in determining boundaries of competence in practice. *Professional Psychology: Research and Practice*, *49*(4), 247–254.

Pituch, K. A., & Stevens, J. P. (2015). *Applied multivariate statistics for the social sciences: Analyses with SAS and IBM's SPSS* (6th ed.). London: Routledge. doi:10.4324/978 1315814919

Okamura, K. H., Jackson, D. S., & Nakamura, B. J. (2019). Therapist and youth predictors of specific practices derived from the evidence-base in community mental health. *Administration and Policy in Mental Health,* doi:10.1007/s10488-019-00942-1

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks: Sage Publications.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R., & Du Toit, M. (2011). HLM statistical software: Version 7. *Computer software*.

Reese, R. J., Norsworthy, L. A., & Rowlands, S. R. (2009a). Does a continuous feedback system improve psychotherapy outcome? *Psychotherapy: Theory, Research, Practice, Training*, *46*(4), 418-431. doi:10.1037/a0017901

Reese, R. J., Usher, E. L., Bowman, D. C., Norsworthy, L. A., Halstead, J. L., Rowlands, S. R., & Chisholm, R. R. (2009b). Using client feedback in psychotherapy training: An analysis of its influence on supervision and counselor self-efficacy. *Training and Education in Professional Psychology, 3*, 157–168. doi:10.1037/a0015673

Riemer, M., Rosof-Williams, J., & Bickman, L. (2005). Theories related to changing clinician practice. *Child and Adolescent Psychiatric Clinics, 14*(2), 241-254.

Sapyta, J., Riemer, M., & Bickman, L. (2005). Feedback to clinicians: Theory, research, and practice. *Journal of Clinical Psychology*, *61*(2), 145-153. doi:10.1002/jclp.20107

Schiefele, A. K., Lutz, W., Barkham, M., Rubel, J., Böhnke, J., Delgadillo, J., ... & Lambert, M. J. (2017). Reliability of therapist effects in practice-based psychotherapy research: A guide for the planning of future studies. *Administration and Policy in Mental Health and Mental Health Services Research*, *44*(5), 598-613.

Shechtman, Z., & Sarig, O. (2016). The effect of client progress feedback on child/adolescent's group-counseling outcomes. *Journal for Specialists in Group Work*, *41*(4), 334-349. doi:10.1080/01933922.2016.1232323

Shimokawa, K., Lambert, M. J., & Smart, D. W. (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology*, *78*(3), 298-311. doi:10.1037/a0019247

Simon, A. E., & Schoendorf, K. C. (2014). Emergency department visits for mental health conditions among U.S. children, 2001-2011. *Clinical Pediatrics*, *53*(14), 1359-1366.

Smith, A. M., & Jensen-Doss, A. (2017). Youth psychotherapy outcomes in usual care and predictors of outcome group membership. *Psychological Services*, *14*(1), 66-76. doi:10.1037/ser0000115

Stein, B. D., Kogan, J. N., Hutchison, S. L., Magee, E. A., & Sorbero, M. J. (2010). Use of outcomes information in child mental health treatment: Results from a pilot study. *Psychiatric Services*, *61*(12), 1211-1216. doi:10.1176/appi.ps.61.12.1211

Torio, C. M., Encinosa, W., Berdahl, T., McCormick, M. C., & Simpson, L. A. (2015). Annual report on health care for children and youth in the United States: National estimates of cost, utilization and expenditures for children with mental health conditions. *Academic Pediatrics*, *15*(1), 19-35. doi:10.1016/j.acap.2014.07.007

Tracey, T. J., Wampold, B. E., Lichtenberg, J. W., & Goodyear, R. K. (2014). Expertise in psychotherapy: An elusive goal? *American Psychologist*, *69*(3), 218.

Treichler, E. B. H., & Spaulding, W. D. (2018). Provider-associated measurement error in routine outcome monitoring in community mental health. *Administration and Policy in Mental Health and Mental Health Services Research*, *45*(5), 780-789. doi:10.1007/s10488-018-0861-2

U.S. Department of Health and Human Services (2017). Justification of estimates for appropriation committees. Retrieved from https://www.cdc.gov/budget/ documents/fy 2017/fy-2017-cdc-congressional-justification.pdf

Walfish, S., McAlister, B., O'Donnell, P., & Lambert, M. J. (2012). An investigation of self-assessment bias in mental health providers. *Psychological Reports*, *110*(2), 639-644. doi:10.2466/02.07.17.PR0.110.2.639-644

Waller, G., & Turner, H. (2016). Therapist drift redux: Why well-meaning clinicians fail to deliver evidence-based therapy, and how to get back on track. *Behaviour Research and Therapy*, *77*, 129-137.

Warren, J. S., Nelson, P. L., Mondragon, S. A., Baldwin, S. A., & Burlingame, G. M. (2010). Youth psychotherapy change trajectories and outcomes in usual care: Community mental health versus managed care settings. *Journal of Consulting and Clinical Psychology, 78*(2), 144-155. doi:10.1037/a0018544

Warren, J. S., Nelson, P. L., Burlingame, G. M., & Mondragon, S. A. (2012). Predicting patient deterioration in youth mental health services: Community mental health vs. managed care settings. *Journal of Clinical Psychology*, *68*(1), 24-40. doi:10.1002/jclp.20831

Weisz, J. R., Chorpita, B. F., Palinkas, L. A., Schoenwald, S. K., Miranda, J., Bearman, S. K., ... & Gray, J. (2012). Testing standard and modular designs for psychotherapy treating depression, anxiety, and conduct problems in youth: A randomized effectiveness trial. *Archives of general psychiatry*, *69*(3), 274-282.

Weisz, J. R., Kuppens, S., Ng, M. Y., Eckshtain, D., Ugueto, A. M., Vaughn-Coaxum, R., & ... Fordwood, S. R. (2017). What five decades of research tells us about the effects of youth psychological therapy: A multilevel meta-analysis and implications for science and practice. *American Psychologist*, *72*(2), 79-117. doi:10.1037/a0040360

Weisz, J. R., Ugueto, A. M., Cheron, D. M., & Herren, J. (2013). Evidence-based youth psychotherapy in the mental health ecosystem. *Journal of Clinical Child & Adolescent Psychology*, *42*(2), 274-286.

Whitcomb, K. E., Woodland, S. C., & Burlingame, G. M. (2018). Do clinicians really use feedback-monitoring systems? A qualitative analysis of 16 group leaders. *Psychotherapy*, *55*(2), 132–143.

Worthen, V. E., & Lambert, M. J. (2007). Outcome oriented supervision: Advantages of adding systematic client tracking to supportive consultations. *Counselling & Psychotherapy Research, 7*(1), 48-53. doi:10.1080/14733140601140873